

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-11

Perfect score: 485

Sequence: 1 LKIDSDSDYVKEGFRAP.....KXAELEKTRADLKKAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 sege, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID      | Description         |
|------------|-------|-------------|--------|------------|---------------------|
| 1          | 485   | 100.0       | 204    | 2 AAW14571 | Aaw14571 Streptococ |
| 2          | 485   | 100.0       | 204    | 7 ABW02605 | Abw02605 Ef1019c p  |
| 3          | 485   | 100.0       | 8991   | 6 ABU08487 | Abu08487 S. pneumo  |
| 4          | 469   | 96.7        | 198    | 7 ABW02615 | Abw02615 Rxl1c pneu |
| 5          | 469   | 96.7        | 315    | 2 AAY04375 | Aay04375 Streptococ |
| 6          | 469   | 96.7        | 619    | 2 AAR63437 | Aar63437 Streptococ |
| 7          | 469   | 96.7        | 619    | 2 AAR87598 | Aar87598 Pneumococ  |
| 8          | 469   | 96.7        | 619    | 2 AAR86911 | Aar86911 Pneumococ  |
| 9          | 469   | 96.7        | 619    | 2 AAY41838 | Aay41838 Streptococ |
| 10         | 469   | 96.7        | 619    | 5 AAE18782 | Aae18782 S. pneumo  |
| 11         | 469   | 96.7        | 619    | 6 ABU45778 | Abu45778 Protein e  |
| 12         | 469   | 96.7        | 619    | 8 ADOS2126 | Ados2126 Streptococ |
| 13         | 469   | 96.7        | 648    | 2 AAW70336 | Aaw70336 Pneumococ  |
| 14         | 469   | 96.7        | 648    | 2 AAW62274 | Aaw62274 Streptococ |
| 15         | 469   | 96.7        | 648    | 2 AAY41837 | Aay41837 Streptococ |
| 16         | 469   | 96.7        | 648    | 2 AAW87879 | Aaw87879 A pneumoc  |
| 17         | 469   | 96.7        | 653    | 2 AAW92456 | Aaw92456 S. pneumo  |
| 18         | 469   | 96.7        | 684    | 2 AAR73912 | Aar73912 Streptococ |
| 19         | 466   | 96.1        | 198    | 2 AAW14581 | Aaw14581 Streptococ |
| 20         | 449   | 92.6        | 653    | 2 AAR27150 | Aar27150 PspA frag  |
| 21         | 446.5 | 92.1        | 289    | 2 AAW62276 | Aaw62276 Streptococ |
| 22         | 446.5 | 92.1        | 289    | 2 AAY41840 | Aay41840 Streptococ |
| 23         | 446.5 | 92.1        | 289    | 2 AAW87910 | Aaw87910 Protein s  |
| 24         | 446.5 | 92.1        | 289    | 2 AAW92458 | Aaw92458 S. pneumo  |
| 25         | 442   | 91.1        | 195    | 2 AAW14591 | Aaw14591 Streptococ |

|    |       |      |      |            |                     |
|----|-------|------|------|------------|---------------------|
| 26 | 442   | 91.1 | 195  | 7 ABW02625 | Abw02625 Wu2c pneu  |
| 27 | 423   | 87.2 | 623  | 6 ABU08494 | Abu08494 Fragment   |
| 28 | 414   | 85.4 | 170  | 7 ABW02614 | Abw02614 Rct135c p  |
| 29 | 414   | 85.4 | 181  | 7 ABW02596 | Abw02596 0922134c   |
| 30 | 414   | 85.4 | 865  | 6 ABU08489 | Abu08489 S. pneumo  |
| 31 | 414   | 85.4 | 929  | 2 AAW14593 | Aaw14593 Streptococ |
| 32 | 414   | 85.4 | 929  | 2 AAY43384 | Aay43384 S. pneumo  |
| 33 | 411   | 84.7 | 188  | 2 AAW14580 | Aaw14580 Streptococ |
| 34 | 411   | 84.7 | 188  | 7 ABW02613 | Abw02613 Rct129c p  |
| 35 | 410   | 84.5 | 1231 | 6 ABU08490 | Abu08490 Fragment   |
| 36 | 402   | 82.9 | 588  | 6 ABU08491 | Abu08491 Coiled co  |
| 37 | 402   | 82.9 | 589  | 2 AAY43392 | Aay43392 PspC alph  |
| 38 | 400   | 82.5 | 204  | 2 AAW14578 | Aaw14578 Streptococ |
| 39 | 400   | 82.5 | 204  | 7 ABW02612 | Abw02612 Rct123c p  |
| 40 | 399.5 | 82.4 | 180  | 2 AAW14562 | Aaw14562 Streptococ |
| 41 | 396.5 | 81.8 | 187  | 2 AAW14579 | Aaw14579 Streptococ |
| 42 | 389   | 80.2 | 206  | 2 AAW14574 | Aaw14574 Streptococ |
| 43 | 389   | 80.2 | 206  | 7 ABW02608 | Abw02608 DB15C pne  |
| 44 | 357.5 | 73.7 | 550  | 8 ADK48356 | Adk48356 Streptococ |
| 45 | 357.5 | 73.7 | 550  | 8 ADR95223 | Adr95223 Novel S.   |

#### ALIGNMENTS

##### RESULT 1

AAW14571

ID AAW14571 standard; protein; 204 AA.

XX

AC AAW14571;

XX

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX

DE Streptococcus pneumoniae PspA central region.

XX

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;

KW bacteraemia; pneumonia.

XX

OS Streptococcus pneumoniae; strain Ef1019.

XX

PN WO9709994-A1.

XX

PD 20-MAR-1997.

XX

PF 16-SEP-1996; 96WO-US014819.

XX

PR 15-SEP-1995; 95US-00529055.

XX

(UABR-) UAB RES FOUND.

PA

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX

WP1; 1997-202002/18.

DR

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used

PT in vaccines for protecting animals against S.pneumoniae infection.

XX

PS Example 6; Fig 13; 296pp; English.

XX

CC This sequence shows the central portion, including the C-terminus of the

CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Ef1019.

CC Comparison of the N-terminal and central regions (AAW14533-57 and

CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can

CC be used to divide the strains into several families based on sequence

CC homologies. PspA polypeptides, or fragments of them, can be used in

CC vaccines to protect animals against S. pneumoniae infection and hence for

CC the prevention of diseases such as otitis media, meningitis, bacteraemia

CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical

CC region and the immediate 5' tip of the coding sequence are likely to be

CC the critical sequences for predicting PspA cross-reactions and vaccine

```

CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 485; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 9.2e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
Db 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99

RESULT 2
ABW02605
ID ABW02605 standard; protein; 204 AA.
XX
AC ABW02605;
XX
XX 12-FEB-2004 (first entry)
XX
DE Ef1019c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
XX US6592876-B1.
XX
PD 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 51; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Ef1019c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
XX Sequence 204 AA;

Query Match 100.0%; Score 485; DB 7; Length 204;
Best Local Similarity 100.0%; Pred. No. 9.2e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
Db 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99

RESULT 3
ABU08487
ID ABU08487 standard; protein; 8991 AA.
XX
AC ABU08487;
XX
XX 24-JUN-2003 (first entry)
XX
DE S. pneumoniae pneumococcal surface protein A (PspA) protein.
XX
KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
KW antibacterial.
XX
OS Streptococcus pneumoniae.
XX
XX Key Location/Qualifiers
FH Misc-difference 1. .8991
FT /note= "All Xaa residues within this sequence are
FT unknown"
XX
XX US6500613-B1.
XX
XX 31-DEC-2002.
XX
XX 16-SEP-1996; 96US-00714741.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 2003-361534/34.
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
PT hybridization probe, comprises pneumococcal surface protein having alpha-
PT helical, proline rich and repeat regions.
XX
XX Disclosure; Col 145-188; 186pp; English.
XX
XX The present invention relates to the isolation of Streptococcus
CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
CC like protein having alpha-helical, proline rich and repeat regions. The
CC PspC and PspA proteins may be used in a vaccine to protect against
CC pneumococcal infections. The polynucleotide sequences encoding PspC and
CC PspA may be used for the expression of the proteins, and as PCR primers
CC or hybridisation probes. The present sequence represents S. pneumoniae
CC PspA protein
XX
XX Sequence 8991 AA;

Query Match 100.0%; Score 485; DB 6; Length 8991;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
Db 5322 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 5381

```



QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99  
 DB 5382 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 5420

## RESULT 4

AW02615  
 ID AEW02615 standard; protein; 198 AA.

XX AC AEW02615;  
 XX DT 12-FEB-2004 (first entry)  
 XX DE Rx1c pneumococcal surface protein A (PspA) central region.

XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 XX OS immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX PN US6592876-B1.

XX PD 15-JUL-2003.

XX PF 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooke-Walter A;

XX DR WPI; 2003-862841/80.

XX PT Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.

XX PS Example 6; SEQ ID NO 61; 121pp; English.

XX CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Rx1c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention

XX SQ Sequence 198 AA;

Query Match 96.7%; Score 469; DB 7; Length 198;  
 Best Local Similarity 96.0%; Pred. No. 2.6e-35;  
 Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVKEGFRAPLOQSELDKAKLSKLELSKIDELDAETAKLEDQKAAE 60

DB 1 LKIDSESESDYAKEGFRAPLOQSKLDKAKLSKLELSKIDELDAETAKLEDQKAAE 60

QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99

DB 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99

## RESULT 5

AAV04375

ID AAY04375 standard; protein; 315 AA.

XX AC AAY04375;

XX DT 23-JUN-1999 (first entry)

XX DE Streptococcus pneumoniae PspA protein sequence.

XX KW Streptococcus pneumoniae; pspA; pneumococcal; surface protein; vaccine;  
 XX OS immunological; infection.

XX OS Streptococcus pneumoniae.

XX OS Synthetic.

XX PN WO9914333-A2.

XX PD 25-MAR-1999.

XX PF 18-SEP-1998; 98WO-US019740.

XX PR 18-SEP-1997; 97US-00932982.

XX PA (INMR ) PASTEUR MERIEUX CONNAUGHT.

XX PI Becker R, Gray M, Pyle D;

XX DR WPI; 1999-229537/19.

XX DR N-PSDB; AAX33124.

XX PT DNA encoding PspA molecule with modified internal translational

XX PT initiation sites.

XX PS Disclosure; Page; 36pp; English.

XX CC The present sequence represents a pneumococcal surface protein A (PspA)  
 CC molecule where internal naturally occurring translational initiation  
 CC sites have been modified or eliminated so that expression of the DNA  
 CC sequence results in a single form of PspA. The PspA nucleotide sequence  
 CC can be used to transform a unicellular host to produce the PspA protein.  
 CC The PspA protein can be used in an immunological composition for treating  
 CC or preventing S. pneumoniae infection especially in a child. Antibodies  
 CC to the PspA protein can also be used to treat S. pneumoniae infection.  
 CC The immunogenic peptides are designed to confer broad protection against  
 CC diverse pneumococcal strains. N.B. The present sequence is not given in  
 CC the specification but is encoded by the sequence given in AAX33124

XX SQ Sequence 315 AA;

Query Match 96.7%; Score 469; DB 2; Length 315;  
 Best Local Similarity 96.0%; Pred. No. 4.7e-35;  
 Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVKEGFRAPLOQSELDKAKLSKLELSKIDELDAETAKLEDQKAAE 60

DB 193 LKIDSESESDYAKEGFRAPLOQSKLDKAKLSKLELSKIDELDAETAKLEDQKAAE 252

QY 61 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 99

DB 253 ENNVEDYFKGLEKTIAAKAELEKTEADLKAVNEPE 291

## RESULT 6

AAE63437

ID AAE63437 standard; protein; 619 AA.

XX AC AAE63437;

XX XX 09-SEP-2004 (revised)

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 19-JUL-1995 (first entry)

XX

DE Pneumococcal surface protein A from S.pneumoniae Rxl.  
 XX  
 KW Pneumococcal surface protein A; PspA; Streptococcus; PCR; pneumococcal;  
 KW primer; protection-eliciting epitope; epitope; vaccine; amplify.  
 XX  
 OS Streptococcus pneumoniae.  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 XX Protein 192..260  
 FT /label= "protein fragment of Claim 1"  
 FT  
 XX EP622081-A2.  
 XX  
 XX 02-NOV-1994.  
 PD  
 XX 19-APR-1994; 94EP-00302767.  
 PF  
 XX 20-APR-1993; 93US-00048896.  
 PR  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, Yother JL, McDaniel LS;  
 XX WPI; 1994-359522/45.  
 DR N-PSDB; AAQ78131.  
 XX  
 XX regions of Pneumococcal surface protein A - derived from the Rxl PspA  
 PT strain, for the preparation of cross-reactive vaccines for the prevention  
 FT of pneumococcal infections.  
 FT  
 XX Disclosure; Page 13-16; 26pp; English.  
 PS  
 XX The amino acid sequence of the novel Pneumococcal surface protein A  
 CC (PspA) from Streptococcus pneumoniae strain Rxl. The gene was PCR  
 CC amplified from S.pneumoniae genomic DNA using the primers AAQ78132-5. The  
 CC gene was used to derive truncated peptide fragments containing protection  
 CC -eliciting epitopes for use in vaccines against pneumococcal diseases.  
 CC The epitopic fragments are derived from amino acids 192-260 and  
 CC optionally contain a further 25 a.a. residues at both the N- and C-  
 CC terminal regions of the peptide. The epitopic peptide fragments may be  
 CC derived from different strains of S.pneumoniae and are homologous to the  
 CC Rxl strain epitope. (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 16-OCT-2003 to standardise OS field)  
 CC  
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key  
 CC  
 XX Sequence 619 AA;  
 SQ  
 Query Match 96.7%; Score 469; DB 2; Length 619;  
 Best Local Similarity 96.0%; Pred. No. 1.1e-34;  
 Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 LKEIDESSEDYVKEGFRAPLQSLDAKQAKLSKLELSKDIDELDAEIAKLEDLKAAE 60  
 Db 223 LKEIDESESDYAKEGFRAPLQSLDAKQAKLSKLELSKDIDELDAEIAKLEDLKAAE 282  
 Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99  
 Db 283 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 321  
 RESULT 7  
 AAR87598  
 ID AAR87598 standard; protein; 619 AA.  
 XX  
 AC AAR87598;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 04-JUL-1996 (first entry)  
 XX  
 DS Pneumococcal surface protein (PspA).

XX PspA; pneumococcal surface protein; truncated; immunoprotective;  
 KW soluble fragment; insertion-duplication mutagenesis.  
 XX  
 OS Streptococcus pneumoniae; strain Rxl.  
 XX  
 XX Key Location/Qualifiers  
 XX Peptide 1..31  
 FT /label= signal\_peptide  
 FT Protein 32..619  
 FT /label= mature\_protein  
 FT Region 32..319  
 FT /label= alpha-helical coiled-coil region  
 FT /note= "contains a seven-residue periodicity"  
 FT  
 FT Region 320..401  
 FT /note= "proline-rich region"  
 FT Region 402..421  
 FT /note= "repeat region"  
 FT Region 422..441  
 FT /note= "repeat region"  
 FT Region 442..461  
 FT /note= "repeat region"  
 FT Region 462..481  
 FT /note= "repeat region"  
 FT Region 482..501  
 FT /note= "repeat region"  
 FT Region 502..521  
 FT /note= "repeat region"  
 FT Region 522..541  
 FT /note= "repeat region"  
 FT Region 542..561  
 FT /note= "repeat region"  
 FT Region 562..581  
 FT /note= "repeat region"  
 FT Region 582..619  
 FT /note= "hydrophobic region starts in last repeat region  
 is potential membrane-spanning region"  
 XX  
 XX US5476929-A.  
 PD 19-DEC-1995.  
 XX  
 XX 03-JUN-1993; 93US-00072070.  
 PF  
 XX 15-FEB-1991; 91US-00656773.  
 PR 12-FEB-1992; 92US-00835698.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 XX  
 XX McDaniel LS, Yother JL, Briles DE;  
 XX WPI; 1996-049021/05.  
 DR N-PSDB; AAT08979.  
 XX  
 XX New pneumococcal surface protein A fragments - comprise proline-rich  
 PT region and/or repeat region, used for detection of Streptococcus  
 PT pneumoniae.  
 XX  
 XX Claim 1; Col 15-20; 23pp; English.  
 PS  
 XX The present sequence is that of PspA (pneumococcal surface protein A)  
 CC encoded by AAT08979. Through the technique of insertion-duplication  
 CC mutagenesis of the PspA gene of the strain Rxl of Streptococcus  
 CC pneumoniae with plasmids contg. cloned fragments of the PspA structural  
 CC gene, it has been possible to produce soluble fragments of PspA that are  
 CC secreted by pneumococci. The method can be used to provide an  
 CC immunoprotective truncated PspA protein. Primers and probes based on the  
 CC present sequence are claimed, and are useful for the detection of (at  
 CC least 32) S. pneumoniae strains. (Updated on 25-MAR-2003 to correct PF  
 CC field.) (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 XX Sequence 619 AA;

|           |   |   |                    |           |             |    |
|-----------|---|---|--------------------|-----------|-------------|----|
|           | Matches   | 95; Conservative  | 3; Mismatches      | 1; Indels | 0; Gaps     | 0  |
| Qy        | 1   | LKEIDESDSDDYVKEGFRAPLQSELDAAKAKLSKLEELSDKIDELDAEIAKLSDQLKAAE              | 60                 |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |
| Dd        | 223   | LKEIDSESEDYAKEGFRAPLQSKLDAKKAKLSKLEELSDKIDELDAEIAKLSDQLKAAE               | 282                |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |
| Qy        | 61  | ENNNVDYFKEGLEKTIAAKKALEKTEADLKAVNEPE                                      | 99                 |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |
| Dd        | 283   | ENNNVDYFKEGLEKTIAAKKALEKTEADLKAVNEPE                                      | 321                |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |
| RESULT 9  |   |   |                    |           |             |    |
| AAAY41838 |   |   |                    |           |             |    |
| ID        | AAAY41838   | standard; protein;  | 619 AA.            |           |             |    |
| XX        | AAAY41838;  |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| DT        | 08-DEC-1999   | (first entry)   |                    |           |             |    |
| DE        |   |   |                    |           |             |    |
| XX        |   | Streptococcus pneumoniae Rx1 PspA protein sequence.                       |                    |           |             |    |
| KW        |   | Streptococcus pneumoniae Rx1; PspA; immunoprotective; vaccine; diagnosis; |                    |           |             |    |
| KW        |   | infection; pneumococcal surface protein A.                                |                    |           |             |    |
| XX        |   | Streptococcus pneumoniae.   |                    |           |             |    |
| OS        |   |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PN        | US5965400-A.  |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PD        | 12-OCT-1999.  |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PF        | 23-MAY-1994;  | 94US-00247491.  |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PR        | 15-FEB-1991;  | 91US-00656773.  |                    |           |             |    |
| PR        | 12-FEB-1992;  | 92US-00835698.  |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PA        | (UABR-) UAB RES FOUND.  |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PI        | Yother JL, Briles DE;   |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| DR        | WPI; 1999-579913/49.  |   |                    |           |             |    |
| DR        | N-PSDB; AAZ25063.   |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PT        | DNA encoding a truncated pneumococcal surface protein A used in the       |   |                    |           |             |    |
| PT        | development of pneumococcal infections.                                   |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| PS        | Claim 1; Fig 3; 27pp; English.  |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| CC        | The present sequence represents Streptococcus pneumoniae Rx1              |   |                    |           |             |    |
| CC        | immunoprotective Pneumococcal surface protein A (PspA). The present       |   |                    |           |             |    |
| CC        | invention also describes a method of forming the immunoprotective         |   |                    |           |             |    |
| CC        | truncated PspA, comprising incorporating a vector comprising the isolated |   |                    |           |             |    |
| CC        | DNA molecule encoding PspA (I), into a bacterium via transformation. (I)  |   |                    |           |             |    |
| CC        | is used to design primers which are capable of detecting a large number   |   |                    |           |             |    |
| CC        | of S. pneumoniae strains, which in turn can be used to diagnose           |   |                    |           |             |    |
| CC        | pneumococcal infection in mammals (e.g. humans), independent of the       |   |                    |           |             |    |
| CC        | strain which has caused it. The PspA protein is used to develop a vaccine |   |                    |           |             |    |
| CC        | against pneumococcal infection comprising, as an immunologically-active   |   |                    |           |             |    |
| CC        | component, a live attenuated or killed bacteria containing a gene coding  |   |                    |           |             |    |
| CC        | for the truncated form of PspA  |   |                    |           |             |    |
| XX        |   |   |                    |           |             |    |
| SQ        | Sequence 619 AA;  |   |                    |           |             |    |
|           | Query Match   | 96.7%;  | Score 469;         | DB 2;     | Length 619; |    |
|           | Best Local Similarity   | 96.0%;  | Pred. NO. 1.1e-34; |           |             |    |
|           | Matches   | 95; Conservative  | 3; Mismatches      | 1; Indels | 0; Gaps     | 0; |
| Qy        | 1   | LKEIDESDSDDYVKEGFRAPLQSELDAAKAKLSKLEELSDKIDELDAEIAKLSDQLKAAE              | 60                 |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |
| Dd        | 223   | LKEIDSESEDYAKEGFRAPLQSKLDAKKAKLSKLEELSDKIDELDAEIAKLSDQLKAAE               | 282                |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |
| Qy        | 61  | ENNNVDYFKEGLEKTIAAKKALEKTEADLKAVNEPE                                      | 99                 |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |
| Dd        | 283   | ENNNVDYFKEGLEKTIAAKKALEKTEADLKAVNEPE                                      | 321                |           |             |    |
|           |   | :     :     :     :     :     :   |                    |           |             |    |

```

RESULT 10
AAE18782
ID AAE18782 standard; protein; 619 AA.
XX
AC AAE18782;
XX
XX 17-MAY-2002 (first entry)
XX
XX S. pneumoniae Rx1 strain pneumococcal surface protein A (PspA).
XX
XX Coiled-coil structural scaffold; heptad repeat; epitope; immune response;
XX cell-mediated immunity; microbial infection; cross-protection; therapy;
XX antimicrobial; vaccine; pneumococcal surface protein A; PspA.
XX
XX Streptococcus pneumoniae.
XX
FH Key Location/Qualifiers
FT Domain 1..314
FT /label= Helical_domain
FT Region 1..303
FT /note= "N-terminal region"
FT Region 38..44
FT /note= "Immunogenic region 3"
FT Region 40..46
FT /note= "Immunogenic region 5"
FT Region 75..80
FT /note= "Immunogenic region 29"
FT Region 82..87
FT /note= "Immunogenic region 52"
FT Region 96..101
FT /note= "Immunogenic region 66"
FT Region 114..119
FT /note= "Immunogenic region 73"
FT Region 130..135
FT /note= "Immunogenic region 78"
FT Region 137..142
FT /note= "Immunogenic region 89"
FT Region 140..145
FT /note= "Immunogenic region 91"
FT Region 152..156
FT /note= "Immunogenic region 95"
FT Domain 153..170
FT /label= Coiled_coil_motif
FT Region 161..164
FT /note= "Immunogenic region 101"
FT Region 166..170
FT /note= "Immunogenic region 116"
FT Region 173..177
FT /note= "Immunogenic region 122"
FT Region 176..180
FT /note= "Immunogenic region 123"
FT Domain 181..198
FT /label= Coiled_coil_motif
FT Region 187..191
FT /note= "Immunogenic region 130"
FT Region 194..198
FT /note= "Immunogenic region 133"
FT Region 215..219
FT /note= "Immunogenic region 140"
FT Region 226..230
FT /note= "Immunogenic region 145"
FT Region 229..232
FT /note= "Immunogenic region 148"
FT Region 260..264
FT /note= "Immunogenic region 166"
FT Region 263..267
FT /note= "Immunogenic region 168"
FT Region 284..287
FT /note= "Immunogenic region 76"
FT Region 286..290
FT /note= "Immunogenic region 179"
FT

Region 294..298
FT /note= "Immunogenic region 182"
FT Region 295..299
FT /note= "Immunogenic region 185"
FT Region 309..313
FT /note= "Immunogenic region 195"
FT Region 322..326
FT /note= "Immunogenic region 206"
XX
XX WO200196368-A2.
XX
XX 20-DEC-2001.
XX
XX 14-JUN-2001; 2001WO-US019168.
XX
XX 14-JUN-2000; 2000US-0211892P.
XX
XX 23-JUN-2000; 2000US-021387P.
XX
XX (CYTO-) CYTOVAX BIOTECHNOLOGIES INC.
XX
XX Houston ME, Hodges RS;
XX
XX WPI; 2002-188298/24.
XX
XX New synthetic peptide derived from naturally occurring microbial and non-
XX microbial protein antigen useful to stimulate and elicit an immune
XX response in an animal.
XX
XX Example 1; Page 90-92; 99pp; English.
XX
XX The invention relates to the uses of coiled-coil structural scaffold to
XX generate structure-specific peptides, including synthetic peptides
XX derived from naturally occurring microbial and non-microbial protein
XX antigens. The structure of the synthetic peptides utilizes a scaffold of
XX heptad repeat units into which epitopes derived from coiled-coil regions
XX of native proteins are spliced. The resulting peptide has a more stable
XX coiled-coil structure, hence improving presentation of the epitopes in a
XX helical conformation. The peptides of the invention are used to stimulate
XX and elicit an immune response in an animal, as vaccine, to treat or
XX prevent microbial infection by several strains and/or species of
XX microorganism, to provide cross-protection to at least one strain and/or
XX species of microorganism and to stimulate antibody production or cell-
XX mediated immunity to the naturally occurring protein. The present
XX sequence is Streptococcus pneumoniae Rx1 strain pneumococcal surface
XX protein A (PspA) which adopts a coiled-coil structure
XX
XX Sequence 619 AA;
XX
Query Match 96.7%; Score 469; DB 5; Length 619;
Best Local Similarity 96.0%; Pred. No. 1.1e-34;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLOSELDAKQKLSKLSBELSKIDELDAEIAKLDELKAAE 60
Db 223 LKEIDSESDYAKEGFRAPLOSKLDAKAKLSKLSBELSKIDELDAEIAKLDELKAAE 282
QY 61 ENNVEDYFKGLEKTIAAKKAELKTEADLKKAENEPE 99
Db 283 ENNVEDYFKGLEKTIAAKKAELKTEADLKKAENEPE 321

RESULT 11
ABU45778
ID ABU45778 standard; protein; 619 AA.
XX
AC ABU45778;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #31305.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX

```



```

AC AAW70336;
XX 18-NOV-1998 (first entry)
XX Pneumococcal surface protein A (PspA).
DE
XX Pneumococcal surface protein A gene; PspA; PspA epitope; vaccine;
KW insertion-duplication mutagenesis; otitis media; meningitis; bacteraemia;
KW pneumonia.
XX Streptococcus pneumoniae.
XX OS
XX FH Key Location/Qualifiers
XX FT Peptide 1..31
XX FT Protein /note= "Signal peptide"
XX FT Region 32..648
XX FT /note= "PspA"
XX FT /note= "alpha-helical coil region representing the
XX FT truncated PspA of the invention"
XX FT Misc-difference 647
XX FT /note= "Encoded by AGG"
XX PN
XX US5804193-A.
XX 08-SEP-1998.
XX 17-MAR-1994; 94US-00214222.
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX (UABR-) UAB RES FOUND.
XX PA
XX Briles DE, Yother JL;
XX WPI; 1998-505588/43.
XX N-PSDB; AAV33264.
XX Truncated pneumococcal surface protein - useful in vaccines against
XX pneumococcal infection.
XX Example 3; Fig 3A-3C; 22pp; English.
XX PS
XX The present sequence represents the Streptococcus pneumoniae Rx1
XX pneumococcal surface protein A (PspA). The invention provides a purified
XX truncated form of PspA, formed by an insertion-duplication mutagenesis
XX technique, comprising of the first 288 N-terminal residues of the mature
XX form of wild-type PspA (AAW70336). The truncated PspA contains
XX immunoprotective epitopes of PspA. The invention claims for a vaccine
XX against pneumococcal infection, comprising live-attenuated or killed S.
XX pneumoniae, containing the gene coding for the truncated PspA protein.
XX The truncated protein, optionally conjugated to a poorly immunogenic or
XX nonimmunogenic molecule, is claimed to be useful in vaccines against
XX pneumococcal infection, especially otitis media, meningitis, bacteraemia
XX and pneumonia
XX CC
XX Sequence 648 AA;
XX Query Match 96.7%; Score 469; DB 2; Length 648;
XX Best Local Similarity 96.0%; Pred. No. 1.2e-34;
XX Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 LKEIDSDSDYKVEGFAPLQSELDKAKQAKLSKLELSKDIDELDAEIAKLEDQKAAE 60
XX | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
XX 223 LKEIDSESDYAKVEGFAPLQSKLDKAKKAKLSKLELSKDIDELDAEIAKLEDQKAAE 282
XX
XX 61 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKA VNEPE 99
XX | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
XX 283 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKA VNEPE 321
XX
XX RESULT 14
XX AAY41837
XX ID AAY41837 standard; protein; 648 AA.

```

---

```

AAW62274
XX ID AAW62274 standard; protein; 648 AA.
XX AC
XX AAW62274;
XX DT 22-SEP-1998 (first entry)
XX DE Streptococcus pneumoniae pspA protein.
XX KW Streptococcus pneumoniae strain Rxi; pspA; immunoprotective; immunogen;
XX KW pneumococcal surface protein A; cholera toxin B subunit; fusion protein;
XX KW antigenic.
XX OS Streptococcus pneumoniae.
XX PN US5753463-A.
XX PD 19-MAY-1998.
XX PF 06-JUN-1995; 95US-00469434.
XX PR 15-FEB-1991; 91US-00656773.
XX PR 12-FEB-1992; 92US-00835698.
XX PR 03-JUN-1993; 93US-00072065.
XX (UABR-) UAB RES FOUND.
XX PA
XX Yother JL, Briles DE;
XX WPI; 1998-311399/27.
XX DR N-PSDB; AAV39470.
XX Truncated pneumococcal surface protein and cholera toxin B sub-unit
XX fusion protein - useful as an immunogen against Streptococcus pneumoniae.
XX Claim 1; Fig 3; 22pp; English.
XX PS
XX The present sequence represents the pneumococcal surface protein A (PspA)
XX protein from Streptococcus pneumoniae. A recombinant DNA molecule has
XX been developed which encodes a fusion protein comprising a truncated form
XX of PspA and cholera toxin B subunit (CTB), where the DNA molecule
XX comprises a nucleotide sequence encoding the truncated PspA linked by an
XX in-frame genetic fusion to a ctxB gene, and where the truncated PspA
XX contains immunoprotective epitopes and up to 90% of the whole PspA
XX protein, except for the cell membrane anchor region. The fusion protein
XX is useful for providing an immunogen to protect neonates and children
XX against S.pneumoniae. Most antigenic proteins of this strain are not
XX immunogenic enough to provide protection. The antigenic epitopes of the
XX fusion protein are directed against capsular polysaccharide antigens of
XX S.pneumoniae, specifically it contains the protective epitopes of PspA.
XX The protein can also be used in solid-phase immunoassay assays, since
XX it is readily bound to supports coated with monosialoganglioside GM1. The
XX fusion protein is more immunogenic against S.pneumoniae than using PspA
XX alone as the immunogen
XX SQ
XX Sequence 648 AA;
XX Query Match 96.7%; Score 469; DB 2; Length 648;
XX Best Local Similarity 96.0%; Pred. No. 1.2e-34;
XX Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 LKEIDSDSDYKVEGFAPLQSELDKAKQAKLSKLELSKDIDELDAEIAKLEDQKAAE 60
XX | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
XX 223 LKEIDSESDYAKVEGFAPLQSKLDKAKKAKLSKLELSKDIDELDAEIAKLEDQKAAE 282
XX
XX 61 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKA VNEPE 99
XX | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
XX 283 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKA VNEPE 321
XX
XX RESULT 15
XX AAY41837
XX ID AAY41837 standard; protein; 648 AA.

```



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-11  
Perfect score: 485  
Sequence: 1 LKEIDSESDYVKGFRAP.....KXAELEKTEADLKXAVNEPE 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCITUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 485   | 100.0       | 99     | 2  | US-08-710-749-10  |
| 2          | 485   | 100.0       | 99     | 4  | US-09-147-875A-11 |
| 3          | 485   | 100.0       | 204    | 4  | US-08-529-055-51  |
| 4          | 485   | 100.0       | 8991   | 4  | US-08-714-741-32  |
| 5          | 469   | 96.7        | 99     | 2  | US-08-710-749-11  |
| 6          | 469   | 96.7        | 198    | 4  | US-08-529-055-61  |
| 7          | 469   | 96.7        | 198    | 1  | US-08-465-746-2   |
| 8          | 469   | 96.7        | 619    | 1  | US-08-214-164-2   |
| 9          | 469   | 96.7        | 619    | 2  | US-08-467-852A-3  |
| 10         | 469   | 96.7        | 619    | 2  | US-08-246-636-2   |
| 11         | 469   | 96.7        | 619    | 2  | US-08-247-491A-3  |
| 12         | 469   | 96.7        | 619    | 2  | US-08-319-795-2   |
| 13         | 469   | 96.7        | 619    | 2  | US-08-468-985-2   |
| 14         | 469   | 96.7        | 619    | 3  | US-08-312-949-2   |
| 15         | 469   | 96.7        | 648    | 1  | US-08-072-070-2   |
| 16         | 469   | 96.7        | 648    | 1  | US-08-469-434-2   |
| 17         | 469   | 96.7        | 648    | 1  | US-08-214-222-2   |
| 18         | 469   | 96.7        | 648    | 2  | US-08-467-852A-2  |
| 19         | 469   | 96.7        | 648    | 2  | US-08-468-718-2   |
| 20         | 469   | 96.7        | 648    | 2  | US-08-247-491A-2  |
| 21         | 469   | 96.7        | 648    | 3  | US-08-446-201-3   |
| 22         | 469   | 96.7        | 695    | 1  | US-08-127-499A-23 |
| 23         | 469   | 96.7        | 695    | 1  | US-08-482-847-23  |
| 24         | 462.5 | 95.4        | 100    | 4  | US-09-147-875A-12 |
| 25         | 457   | 94.2        | 288    | 3  | US-08-312-949-4   |
| 26         | 457   | 94.2        | 288    | 3  | US-08-446-201-4   |
| 27         | 446.5 | 92.1        | 289    | 1  | US-08-072-070-4   |

|    |       |      |      |   |                   |                    |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 446.5 | 92.1 | 289  | 1 | US-08-469-434-4   | Sequence 4, Appli  |
| 29 | 446.5 | 92.1 | 289  | 1 | US-08-214-222-4   | Sequence 4, Appli  |
| 30 | 446.5 | 92.1 | 289  | 2 | US-08-467-852A-5  | Sequence 5, Appli  |
| 31 | 446.5 | 92.1 | 289  | 2 | US-08-468-718-4   | Sequence 4, Appli  |
| 32 | 446.5 | 92.1 | 289  | 2 | US-08-247-491A-5  | Sequence 5, Appli  |
| 33 | 446   | 92.0 | 99   | 4 | US-09-147-875A-13 | Sequence 13, Appli |
| 34 | 442   | 91.1 | 99   | 2 | US-08-710-749-12  | Sequence 12, Appli |
| 35 | 442   | 91.1 | 195  | 4 | US-08-529-055-71  | Sequence 71, Appli |
| 36 | 423   | 87.2 | 623  | 4 | US-08-714-741-47  | Sequence 47, Appli |
| 37 | 414   | 85.4 | 170  | 4 | US-08-529-055-60  | Sequence 60, Appli |
| 38 | 414   | 85.4 | 181  | 4 | US-08-529-055-42  | Sequence 42, Appli |
| 39 | 414   | 85.4 | 864  | 4 | US-08-714-741-40  | Sequence 40, Appli |
| 40 | 413   | 85.2 | 99   | 2 | US-08-710-749-17  | Sequence 17, Appli |
| 41 | 412.5 | 85.1 | 100  | 4 | US-09-147-875A-10 | Sequence 10, Appli |
| 42 | 411   | 84.7 | 99   | 4 | US-09-147-875A-16 | Sequence 16, Appli |
| 43 | 411   | 84.7 | 188  | 4 | US-08-529-055-59  | Sequence 59, Appli |
| 44 | 410   | 84.5 | 1231 | 4 | US-08-714-741-41  | Sequence 41, Appli |
| 45 | 402   | 82.9 | 141  | 4 | US-09-286-981B-2  | Sequence 2, Appli  |

ALIGNMENTS

RESULT 1  
US-08-710-749-10  
; Sequence 10, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-10

Query Match 100.0%; Score 485; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.1e-37;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LKEIDSESDYVKGFRAPLOSELDKAKLSKLEELSDKIDELDAETAKLEDQKAAE 60  
Db 1 LKEIDSESDYVKGFRAPLOSELDKAKLSKLEELSDKIDELDAETAKLEDQKAAE 60

QY 61 ENNVEDYFKEGLEKTI AAKKAELEKTEADLKKAVNEPE 99

Db 61 ENNVEDYFKEGLEKTI AAKKAELEKTEADLKKAVNEPE 99

```

RESULT 2
US-09-147-875A-11
; Sequence 11, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-11

```

```

Query Match      100.0%; Score 485; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 2.1e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LKEIDSDSDYVKEGPRAPLQSELDAAQAKLSKLELSDKIDELDAEIAKLELDQAKAE 60
Db      1 LKEIDSDSDYVKEGPRAPLQSELDAAQAKLSKLELSDKIDELDAEIAKLELDQAKAE 60

Qy      61 ENNNVEDYFKEGLEKTTAAKAAELEKTEADLKKAVNEPE 99
Db      61 ENNNVEDYFKEGLEKTTAAKAAELEKTEADLKKAVNEPE 99

```

RESULT 3  
US-08-529-055-51  
Sequence 51, Application US/0829055  
Patent No. 6592876  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yohet, Janet  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: Pneumococcal Genes, Portions  
TITLE OF INVENTION: Thereof, Expression Products  
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
TITLE OF INVENTION: Portions and Products  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400

Query Match 100.0%; Score 485; DB 4; Length 8991;  
Best Local Similarity 100.0%; Pred. No. 4.5e-35;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60  
DB 5322 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 5381  
QY 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99  
DB 5382 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 5420

## RESULT 5

US-08-710-749-11  
; Sequence 11, Application US/08710749  
; Patent No. 5955089

## GENERAL INFORMATION:

; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712

## INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid

US-08-710-749-11

Query Match 96.7%; Score 469; DB 2; Length 99;  
Best Local Similarity 96.0%; Pred. No. 6.2e-36;  
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60  
DB 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60  
QY 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99  
DB 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99

## RESULT 6

US-08-529-055-61  
; Sequence 61, Application US/08529055  
; Patent No. 6592876

## GENERAL INFORMATION:

; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-529-055-61

Query Match 96.7%; Score 469; DB 4; Length 198;  
Best Local Similarity 96.0%; Pred. No. 1.4e-35;  
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60  
DB 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60  
QY 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99  
DB 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 99

## RESULT 7

US-08-465-746-2  
; Sequence 2, Application US/08465746  
; Patent No. 5679768

## GENERAL INFORMATION:

; APPLICANT: Briles, David E.  
; APPLICANT: Yother, Janet L  
; APPLICANT: McDaniel, Larry S  
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEIN A  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd  
; STREET: Suite 1203, 2001 Jefferson Davis Highway  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.

```
;
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/465,746
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/048,896
; FILING DATE:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-746-2

Query Match 96.7%; Score 469; DB 1; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKAKKLSKLELSKIDELDAEIAKLEDLQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEDLQKAAE 282
QY 61 ENNVEDYFKGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
Db 283 ENNVEDYFKGLEKTTIAAKKAELEKTEADLKKA VNEPE 321

RESULT 8
US-08-214-164-2
; Sequence 2, Application US/08214164
; Patent No. 5728387
; GENERAL INFORMATION:
; APPLICANT: BRILES, DAVID E.
; APPLICANT: YOTHER, JANET L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/214,164
; FILING DATE: 17-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:

; NAME: Berkstresser, Jerry W.
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 6102-137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0813
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-214-164-2

Query Match 96.7%; Score 469; DB 1; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKAKKLSKLELSKIDELDAEIAKLEDLQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEDLQKAAE 282
QY 61 ENNVEDYFKGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
Db 283 ENNVEDYFKGLEKTTIAAKKAELEKTEADLKKA VNEPE 321

RESULT 9
US-08-467-852A-3
; Sequence 3, Application US/08467852A
; Patent No. 5856170
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTHER, Janet L.
; APPLICANT: MCDANIEL, Larry S.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,852A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-467-852A-3

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

|    |     |   |     |
|----|-----|---|-----|
| Qy | 1   | LKSIDSDSDYVKGGRFAPQSELDADKQAKLSKLEELSDKIDELDAEIAKLBDQLKAAE        | 60  |
|    |     | :       :       :       :       :       :       :       :       : |     |
| Db | 223 | LKSIDSESDYAKGGFRAPQSKLDADKQAKLSKLEELSDKIDELDAEIAKLBDQLKAAE        | 282 |
| Qy | 61  | ENNNVDYFKEGLEKTTIAAKKAELEKTEADLKAVNPE                             | 99  |
| Db | 283 | ENNNVDYFKEGLEKTTIAAKKAELEKTEADLKAVNPE                             | 321 |

RESULT 10  
 US-08-246-636-2  
 ; Sequence 2, Application US/08246636  
 ; Patent No. 5965141  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E.  
 ; APPLICANT: Yother, Janet L.  
 ; APPLICANT: Mcdaniel, Larry S.  
 ; APPLICANT: Wu, Hong-Yin  
 ; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE  
 ; TITLE OF INVENTION: PROTEIN A  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Shoemaker and Mattare, Ltd  
 ; STREET: Suite 1203, 2001 Jefferson Davis Highway  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22202-0286  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/246,636  
 ; FILING DATE: 20-MAY-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/656,773  
 ; FILING DATE: 15-FEB-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/835,698  
 ; FILING DATE: 12-FEB-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/048,896  
 ; FILING DATE: 20-APR-1993  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 415-0810  
 ; TELEFAX: (703) 415-0813  
 ; TELEX: LUKPAT WASHINGTON  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 619 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-246-636-2

|                       |       |                    |               |                   |
|-----------------------|-------|--------------------|---------------|-------------------|
| Query Match           | 96.7% | Score 469;         | DB 2;         | Length 619;       |
| Best Local Similarity | 96.0% | Pred. No. 5.4e-35; |               |                   |
| Matches               | 95;   | Conservative 3;    | Mismatches 1; | Indels 0; Gaps 0; |

  

|    |     |   |     |
|----|-----|---|-----|
| Qy | 1   | LKEIDSDSDYVVEGFRAPLQSELDAAQAKLSKLELSDKIDELDAEIAKLEPQLKAAE | 60  |
|    |     | :   |     |
|    |     | :   |     |
| Db | 223 | LKEIDSESDYAKEGFRAPLQSKLDAAKAKLSKLELSDKIDELDAEIAKLEPQLKAAE | 282 |

  

|    |     |  |     |
|----|-----|--|-----|
| Qy | 61  | ENNVDYFFKEGLEKTTAAKAAELEKTEADLKKAVNEPE | 99  |
|    |     | :                                      |     |
|    |     | :                                      |     |
| Db | 283 | ENNVDYFFKEGLEKTTAAKAAELEKTEADLKKAVNEPE | 321 |

RESULT 11

```

US-08-247-491A-3
; Sequence 3, Application US/08247491A
; Patent No. 5965400
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTHER, Janet L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,491A
; FILING DATE: 23-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-247-491A-3

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0

Qy 1 LKIDSDSDYVKEGFRAPLQSELDKAKLKLSELSDKIDELDAETAKI
Db 223 LKIDSESDYAKGFRAPLQSKLDKAKLKLSELSDKIDELDAETAKI
Qy 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEP 99
Db 283 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEP 321

RESULT 12
US-08-319-795-2
; Sequence 2, Application US/08319795
; Patent No. 5980909
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
; TITLE OF INVENTION: Protein A
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheomaker and Mattare, Ltd.
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:

```

```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA: US/08/319,795
/ FILING DATE:
/ APPLICATION NUMBER: US/08/319,795
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/246,636
/ FILING DATE: 20-MAY-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/048,896
/ FILING DATE: 20-APR-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/835,698
/ FILING DATE: 12-FEB-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/656,773
/ FILING DATE: 15-FEB-1991
/ APPLICATION NUMBER: US 08/048,896
/ FILING DATE: 20-APR-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/835,698
/ FILING DATE: 12-FEB-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/656,773
/ FILING DATE: 15-FEB-1991
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 415-0810
/ TELEFAX: (703) 415-0813
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 619 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-319-795-2

Query Match          96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESEDYAKEGFRAPLQSKLDKAKLSKLELSKIDELDAEIAKLEDLKAAE 282

Qy 61 ENNVEDYFKEGLEKTTAAKKALEKTEADLKAVNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKEGLEKTTAAKKALEKTEADLKAVNEPE 321

RESULT 13
US-08-468-985-2
/ Sequence 2, Application US/08468985
/ Patent No. 5997882
/ GENERAL INFORMATION:
/ APPLICANT: Briles, David E.
/ APPLICANT: Yother, Janet L.
/ APPLICANT: McDaniel, Larry S
/ TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
/ TITLE OF INVENTION: Protein A
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sheomaker and Mattare, Ltd.
/ STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
/ STREET: Davis Highway
/ CITY: Arlington
/ STATE: Virginia
/ COUNTRY: U.S.A.
/ ZIP: 22202-0286
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/468,985
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:

Query Match          96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 5.4e-35;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESEDYAKEGFRAPLQSKLDKAKLSKLELSKIDELDAEIAKLEDLKAAE 282

Qy 61 ENNVEDYFKEGLEKTTAAKKALEKTEADLKAVNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKEGLEKTTAAKKALEKTEADLKAVNEPE 321

RESULT 14
US-08-112-949-2
/ Sequence 2, Application US/08312949
/ Patent No. 6027734
/ GENERAL INFORMATION:
/ APPLICANT: Briles, David E.
/ APPLICANT: Wu, Hong-Yin
/ TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
/ TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Curtiss, Morris & Safford, P.C.
/ STREET: 530 Fifth Avenue
/ CITY: New York
/ STATE: New York
/ COUNTRY: United States of America
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/312,949
/ FILING DATE: 30-SEP-1994
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Frommer, William S.
/ REGISTRATION NUMBER: 25,506
/ REFERENCE/DOCKET NUMBER: 454312-2049
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 840-3333
/ TELEFAX: (212) 840-0712
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 619 amino acids
```

;  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-312-949-2

Query Match 96.7%; Score 469; DB 3; Length 619;  
Best Local Similarity 96.0%; Pred. No. 5.4e-35;  
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 223 LKIDSESDYAKGFRAPLQSKLDAKKAALKLELSKIDELDAEIAKLEDLQKAAE 282  
QY 61 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKAVNEPE 99  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 283 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKAVNEPE 321

## RESULT 15

US-08-072-070-2  
; Sequence 2, Application US/08072070  
; Patent No: 5476929  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: Yother, Janet L  
; APPLICANT: McDaniel, Larry S  
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd  
; STREET: Suite 1203, 2001 Jefferson Davis Highway  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/072,070  
; FILING DATE: 19930603  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/835,698  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/656,773  
; FILING DATE: 15-FEB-1991  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 415-0810  
; TELEFAX: (703) 521-0378  
; TELEX: LUKPAT WASHINGTON  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 648 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-072-070-2

Query Match 96.7%; Score 469; DB 1; Length 648;  
Best Local Similarity 96.0%; Pred. No. 5.7e-35;  
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 223 LKIDSESDYAKGFRAPLQSKLDAKKAALKLELSKIDELDAEIAKLEDLQKAAE 282  
QY 61 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKAVNEPE 99

DB 283 ENNVEDYFKGLEKTIAAKKAELEKTEADLKKAVNEPE 321  
Search completed: June 21, 2005, 10:25:19  
Job time : 19.4867 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-11

Perfect score: 485

Sequence: 1 LKEIDSDSEDYVKEGFRAP.....KKALEKTEADLKKA VNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 485   | 100.0       | 99     | 15 | US-10-674-755-11     |
| 2          | 485   | 100.0       | 204    | 15 | US-10-299-636-66     |
| 3          | 469   | 96.7        | 198    | 15 | US-10-299-636-76     |
| 4          | 469   | 96.7        | 354    | 15 | US-10-299-636-105    |
| 5          | 469   | 96.7        | 588    | 10 | US-10-299-636-96     |
| 6          | 469   | 96.7        | 619    | 10 | US-09-882-774-1      |
| 7          | 469   | 96.7        | 619    | 16 | US-10-282-122A-73702 |
| 8          | 469   | 96.7        | 619    | 16 | US-10-414-532-72     |
| 9          | 462.5 | 95.4        | 100    | 15 | US-10-674-755-12     |
| 10         | 446   | 92.0        | 99     | 15 | US-10-674-755-13     |
| 11         | 442   | 91.1        | 195    | 15 | US-10-299-636-86     |
|            |       |             |        |    | Sequence 11, Appl    |
|            |       |             |        |    | Sequence 66, Appl    |
|            |       |             |        |    | Sequence 76, Appl    |
|            |       |             |        |    | Sequence 105, Appl   |
|            |       |             |        |    | Sequence 96, Appl    |
|            |       |             |        |    | Sequence 1, Appl     |
|            |       |             |        |    | Sequence 73702, A    |
|            |       |             |        |    | Sequence 72, Appl    |
|            |       |             |        |    | Sequence 12, Appl    |
|            |       |             |        |    | Sequence 13, Appl    |
|            |       |             |        |    | Sequence 86, Appl    |

|    |       |      |     |    |                    |                   |
|----|-------|------|-----|----|--------------------|-------------------|
| 12 | 414   | 85.4 | 170 | 15 | US-10-299-636-75   | Sequence 75, Appl |
| 13 | 414   | 85.4 | 181 | 15 | US-10-299-636-57   | Sequence 57, Appl |
| 14 | 414   | 85.4 | 643 | 15 | US-10-299-636-95   | Sequence 95, Appl |
| 15 | 414   | 85.4 | 670 | 9  | US-09-748-875-63   | Sequence 63, Appl |
| 16 | 414   | 85.4 | 670 | 10 | US-09-298-523B-63  | Sequence 63, Appl |
| 17 | 414   | 85.4 | 690 | 9  | US-09-748-875-61   | Sequence 61, Appl |
| 18 | 414   | 85.4 | 690 | 10 | US-09-298-523B-61  | Sequence 61, Appl |
| 19 | 414   | 85.4 | 691 | 9  | US-09-748-875-1    | Sequence 1, Appl  |
| 20 | 414   | 85.4 | 691 | 10 | US-09-298-523B-1   | Sequence 1, Appl  |
| 21 | 414   | 85.4 | 701 | 9  | US-09-748-875-62   | Sequence 62, Appl |
| 22 | 414   | 85.4 | 701 | 10 | US-09-298-523B-62  | Sequence 62, Appl |
| 23 | 414   | 85.4 | 707 | 9  | US-09-748-875-2    | Sequence 2, Appl  |
| 24 | 414   | 85.4 | 707 | 10 | US-09-298-523B-2   | Sequence 2, Appl  |
| 25 | 414   | 85.4 | 711 | 9  | US-09-748-875-3    | Sequence 3, Appl  |
| 26 | 414   | 85.4 | 711 | 10 | US-09-298-523B-3   | Sequence 3, Appl  |
| 27 | 414   | 85.4 | 739 | 17 | US-10-732-923-3294 | Sequence 3294, Ap |
| 28 | 414   | 85.4 | 929 | 9  | US-09-748-875-60   | Sequence 60, Appl |
| 29 | 414   | 85.4 | 929 | 10 | US-09-298-523B-60  | Sequence 60, Appl |
| 30 | 414   | 85.4 | 929 | 15 | US-10-299-636-94   | Sequence 94, Appl |
| 31 | 412.5 | 85.1 | 100 | 15 | US-10-674-755-10   | Sequence 10, Appl |
| 32 | 411   | 84.7 | 99  | 15 | US-10-674-755-16   | Sequence 16, Appl |
| 33 | 411   | 84.7 | 188 | 15 | US-10-299-636-74   | Sequence 74, Appl |
| 34 | 402   | 82.9 | 141 | 14 | US-10-254-995-2    | Sequence 2, Appl  |
| 35 | 402   | 82.9 | 589 | 9  | US-09-748-875-14   | Sequence 14, Appl |
| 36 | 402   | 82.9 | 589 | 10 | US-09-298-523B-14  | Sequence 14, Appl |
| 37 | 402   | 82.9 | 589 | 15 | US-10-299-636-97   | Sequence 97, Appl |
| 38 | 400   | 82.5 | 204 | 15 | US-10-299-636-73   | Sequence 73, Appl |
| 39 | 393   | 81.0 | 99  | 15 | US-10-674-755-15   | Sequence 15, Appl |
| 40 | 389   | 80.2 | 206 | 15 | US-10-299-636-69   | Sequence 69, Appl |
| 41 | 386   | 79.6 | 99  | 15 | US-10-674-755-14   | Sequence 14, Appl |
| 42 | 352.5 | 72.7 | 100 | 15 | US-10-674-755-2    | Sequence 2, Appl  |
| 43 | 348.5 | 71.9 | 100 | 15 | US-10-674-755-3    | Sequence 3, Appl  |
| 44 | 343.5 | 70.8 | 194 | 15 | US-10-299-636-79   | Sequence 79, Appl |
| 45 | 339.5 | 70.0 | 98  | 15 | US-10-674-755-1    | Sequence 1, Appl  |

ALIGNMENTS

RESULT 1

US-10-674-755-11  
; Sequence 11, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-11

Query Match 100.0%; Score 485; DB 15; Length 99;  
Best Local Similarity 100.0%; Pred. No. 3.1e-32;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSDSEDYVKEGFRAPLOSELDAKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60  
Db 1 LKEIDSDSEDYVKEGFRAPLOSELDAKQAKLSKLELSDKIDELDAETAKLEDQKAAE 60

QY 61 ENNVNEDYFKEGLEKTIAAKKALEKTEADLKKA VNEPE 99  
Db 61 ENNVNEDYFKEGLEKTIAAKKALEKTEADLKKA VNEPE 99

RESULT 2

```
US-10-299-636-66
; Sequence 66, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-66

Query Match      100.0%; Score 485; DB 15; Length 204;
Best Local Similarity 100.0%; Pred. No. 7e-32;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAE 60
Db 1 LKEIDSESDYVKEGFRAPLQSKLDAAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAE 60

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99

US-10-299-636-76
; Sequence 76, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-76

Query Match      96.7%; Score 469; DB 15; Length 198;
Best Local Similarity 96.0%; Pred. No. 1.4e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 LKEIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAE 60
Db 1 LKEIDSESDYVKEGFRAPLQSKLDAAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAE 60

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99

RESULT 4
US-10-299-636-105
; Sequence 105, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-105

Query Match      96.7%; Score 469; DB 15; Length 354;
Best Local Similarity 96.0%; Pred. No. 2.6e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYVKEGFRAPLQSELDAAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAE 60
Db 175 LKEIDSESDYVKEGFRAPLQSKLDAAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAE 234

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
Db 235 ENNVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 273

RESULT 5
US-10-299-636-96
; Sequence 96, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
```

```
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636--96

Query Match          96.7%; Score 469; DB 15; Length 588;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQKAAE 251
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 290
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; TITLE OF INVENTION: Structure-Specific Peptides
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

Query Match          96.7%; Score 469; DB 10; Length 619;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQKAAE 282
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 321
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-10-282-122A-73702
; Sequence 73702, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Hasebeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636--96

; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73702
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73702

Query Match          96.7%; Score 469; DB 15; Length 619;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQKAAE 282
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 ENNVEDYFKEGLEKTTAAKKALEKTEADLKKA VNEPE 321
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
US-10-414-532-72
; Sequence 72, Application US/10414532
; Publication No. US20040101531A1
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 72
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-414-532-72

Query Match          96.7%; Score 469; DB 16; Length 619;
Best Local Similarity 96.0%; Pred. No. 4.9e-30;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 LKEIDSDSDYKVEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDSESEDYAKEGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 321

RESULT 9
US-10-674-755-12
; Sequence 12, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-12

Query Match 95.4%; Score 462.5; DB 15; Length 100;
Best Local Similarity 96.0%; Pred. No. 2.1e-30;
Matches 96; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LKEIDSDSDYKVEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQ-KAA 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSESEDYAKEGFRAPLQSKLDKAKKLSKLELSKIDELDAEIAKLEDLQKAA 60

Qy 60 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 100

RESULT 10
US-10-674-755-13
; Sequence 13, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-13

Query Match 92.0%; Score 446; DB 15; Length 99;
Best Local Similarity 91.9%; Pred. No. 4.6e-29;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKVEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSESEDYAKEGFRAPLHSHKLDKAKKLSKLELSKIDELDAEIAKLEDLQKAVE 60

Qy 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
Db 61 ENNVEDYSTEGLEKTTIAAKKTELEKTEADLKKA VNEPE 99

RESULT 11
US-10-299-636-86
; Sequence 86, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-86

Query Match 91.1%; Score 442; DB 15; Length 195;
Best Local Similarity 90.9%; Pred. No. 2.1e-28;
Matches 90; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYKVEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSESEDYAKEGFRAPLHSHKLDKAKKLSKLELSKIDELDAEIAKLEDLQKAVE 60

Qy 61 ENNVEDYFKEGLEKTTIAAKKAELEKTEADLKKA VNEPE 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ENNVEDYSTEGLEKTTIAAKKTELEKTEADLKKA VNEPE 99

RESULT 12
US-10-299-636-75
; Sequence 75, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75
```

Search completed: June 21, 2005, 11:18:33  
Job time : 64.2388 secs

**This Page Blank (uspto)**

J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:M74122; NID:g153840; PIDN:AAA2701f  
A:Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBI:P:75636)  
R:Talkington, D.F.; Crummins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.





QY 57 KAAEENNV-----EDYFEGLEKTIKAAKAELEKTEADLKAAVNE 97  
 Db 69 QLAEDSNLNENNTYKQND-----LEQLEDSKAEADK 108

RESULT 6

T34418  
 hypothetical protein F12F3.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34418  
 R:Pullton, B.; Wohldmann, P.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: The sequence of C. elegans cosmid F12F3.  
 A:Reference number: Z21521  
 A:Accession: T34418  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-3488 <FUL>  
 A:Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3  
 A:Experimental source: strain Bristol N2; clone F12F3  
 C:Genetics:  
 A:Gene: CESP:F12F3.3  
 A:Map position: 5  
 A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 22.8%; Score 110.5; DB 2; Length 3488;  
 Best Local Similarity 34.4%; Pred. No. 12;  
 Matches 43; Conservative 21; Mismatches 26; Indels 35; Gaps 8;

QY 2 KEIDES---DSEDYVGEFRAPLQSELDKQAKLSKL-----EELSDKIDELDARI 49  
 Db 1009 KETDEKLDAEIAAKTKQEADEKSLDA-QEKIKKVSDDAARKEKELNDKL-KLESRI 1066

QY 50 A-----KLEDO-----LKAEEENNVEDYK---EGLEKTIKAAKAELEKTEA 89  
 Db 1067 ATKASADKLEEQAKAAEVAEAKKQEKDEQLKLDTEAASKAAAEKLELEK-QA 1125

QY 90 DLKKA 94  
 Db 1126 QIKKA 1130

RESULT 7

A28313  
 glued protein - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
 C:Accession: A28313  
 R:Swarcop, A.; Swarcop, M.; Garen, A.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987  
 A:Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued ge  
 A:Reference number: A28313; MUID:87317680; PMID:2819881  
 A:Accession: A28313  
 A:Molecule type: DNA; mRNA  
 A:Residues: 1-1319 <SWA>  
 A:Cross-references: UNIPROT:PI3496  
 A>Note: the authors' translation is inconsistent with the nucleotide sequence in the reg  
 C:Genetics:  
 A:Gene: FlyBase:Gl  
 A:Cross-references: FlyBase:FBgn0001108  
 C:Introns: 18/2; 479/3  
 C:Keywords: cytoskeleton; glycoprotein  
 F:397,590,771,889,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #e

Query Match 22.2%; Score 107.5; DB 2; Length 1319;  
 Best Local Similarity 32.5%; Pred. No. 7.4;  
 Matches 37; Conservative 24; Mismatches 30; Indels 23; Gaps 5;

QY 1 LKEIDSDSEDYVKEGFRAPLQSELDKQAKLSKL-----EELSDKIDELDARIKLEDO 56  
 Db 429 LRDLSDHKHDQK-----LSKELEMKRSEVTELETKESAKIDELEIAVDLQEQV 482

QY 57 KA-----ASENNNVEDYFKEGLEKTIKAAKAELEKTEADLKAAVNEPE 99  
 Db 483 DAALGAEEVVEQAEKQMELEDKVL-LSEETIAQLEA-LEEVHEQLVESNHSLE 534

RESULT 8

S70531  
 bkb2.11 protein precursor - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S70531  
 R:Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard  
 Mol. Microbiol. 18, 507-520, 1995  
 A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer  
 A:Reference number: S70531; MUID:96342380; PMID:8748034  
 A:Accession: S70531  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-233 <AKI>  
 A:Cross-references: UNIPROT:Q44739; EMBL:U30617; NID:g3309515; PIDN:AAC46421.1; PID:g119:  
 C:Superfamily: outer surface protein F ospF  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-233/Product: bkb2.11 protein #status predicted <MAT>

Query Match 22.1%; Score 107; DB 2; Length 233;  
 Best Local Similarity 29.8%; Pred. No. 1.4;  
 Matches 37; Conservative 25; Mismatches 34; Indels 28; Gaps 6;

QY 1 LKEIDES--DSDYK-----EGFRAPLQ---SELDKQAK--LSKLEELSDKI 42  
 Db 29 LKNSQNLESSEQNKKTEQETIKQVEGFLEITKDLSEKDEKTEKQIQELNKKI 88

QY 43 DELDAEIAKLE-----DOLKAAEENNVEDYFKEGLEKTIKAAKAELEKTEADLKK 93  
 Db 89 EKLDKSKTSIETYSEVEEKINKIKELKKGLEDKFKK-LEESLAKKKGKCKALQEAQ 147

QY 94 AVNE 97  
 Db 148 KFEF 151

RESULT 9

H69378  
 conserved hypothetical protein AF1032 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Accession: H69378  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.N  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: H69378  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-886 <KLE>  
 A:Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB90211  
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.1%; Score 107; DB 2; Length 886;  
 Best Local Similarity 25.9%; Pred. No. 5.4;  
 Matches 35; Conservative 25; Mismatches 29; Indels 46; Gaps 5;

QY 1 LKEIDESDSEDYVKEGFRAPLQSELDKQAKLSKLEELSDKIDELDAEI----- 49  
 Db 303 LRDEVKREG-DLTREA--AGIQALKKABEDNSKLEETIKRIEELERLEPERFEKSHRLLE 359

QY 50 -----AKLEDO-----LKAEEENNVEDYFKEGLEKTIKAAKKA 82  
 Db 360 TLKPKMQRQGIKAKLEKLNLPDKVKRWYDLLSKAKEEBEKITEKLRK-----LIAKKS 414

QY 83 ELEKTEADLKAVNE 97  
|:  
Db 415 SLKTRGAQLKAAVEE 429

RESULT 10  
S70532  
outer surface protein F precursor - Lyme disease spirochete  
C;Species: Borrelia burgdorferi (Lyme disease spirochete)  
C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S70532  
R;Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Morgan  
Mol.: Microbiol. 18, 507-520, 1995  
A;Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer  
A;Reference number: S70531; PMID:96342380; PMID:8748034  
A;Accession: S70532  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-229 <AKI>  
A;Cross-references: UNIPROT:Q44735; EMBL:U19754; NID:g3318660; PIDN:AAC26147.1; PID:g9896  
C;Genetics:  
A;Gene: ospF  
C;Superfamily: outer surface protein F ospF  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-229/Product: outer surface protein F #status predicted <MAT>

Query Match 21.6% Score 105; DB 2; Length 229;  
Best Local Similarity 31.6%; Pred. No. 1.9;  
Matches 37; Conservative 21; Mismatches 33; Indels 26; Gaps 6;

QY 5 DESDSEDYVK-----EGFRAPLQ----SELDAQAQLSKLELSKDIDELDAETA 50  
| | | | |  
Db 35 DLSSSQNVKKTRQEIKKQVGEFLLETLDLNKLDTKEIE-KRIQELKEKIEKLEAKKT 93  
| | | | |  
QY 51 KL-----EQDKAAEE---NNNVEDYFKEGLEKTTAAKKALEKTEADLKAVNE 97  
| | | | |  
Db 94 SLKYTYEYEKLLQIREKLGKAGDLKGLK-GLEDSLKXKKEERKKALEDAAKKFFEE 149

RESULT 11  
T05409  
hypothetical protein F10M6.170 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T05409  
R;Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; N  
submitted to the Protein Sequence Database, February 1998  
A;Reference number: Z15414  
A;Accession: T05409  
A;Molecule type: DNA  
A;Residues: 1-764 <BEV>  
A;Cross-references: UNIPROT:O49371; EMBL:AL021811  
A;Experimental source: cultivar Columbia; BAC clone F10M6  
C;Genetics:  
A;Map position: 4  
A;Note: F10M6.170

Query Match 21.4% Score 104; DB 2; Length 764;  
Best Local Similarity 31.6%; Pred. No. 7.4;  
Matches 36; Conservative 22; Mismatches 38; Indels 18; Gaps 4;

QY 2 KEIDESDESYVKEGFRAPLOSELDAQAQLSKL-BELS KDIDELAEINAKLEDQKAAE 60  
| | | : : : : :  
Db 163 REIELKHKLRRDEERAALQSLLTKEEELEKMRQEIANNRSKVSMAISEFEFSKQLLS 222  
| | | : : : : :  
QY 61 ENNVN-----EDYF-----KEGLEKTTAAKKAEEK---TEADLKAVNE 97  
| | | : : : : :  
Db 223 KANEVVKRQGEIYALORALEEEEELEISKATYKLEOEKLRTEANLKKQTTEE 276

RESULT 12  
T26963

```

A;Molecule type: DNA
A;Residues: 1-852 <ARN>
A;Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:94982196; PIDN:AAD3670
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1636
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match      20.9%; Score 101.5; DB 2; Length 852;
Best Local Similarity 29.2%; Pred. No. 12;
Matches 26; Conservative 23; Mismatches 29; Indels 11; Gaps 3;

QY      6 ESDSEDYVKEGFRAPLQSELDKQAKLSKLE-----LSDKIDELDAEIAKLEQDKAAEE 61
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      506 EKTEELHRLGYSEDLLEKLDKRRKKRKIEEERHSISQKITAADVQISQIENQLK--EI 563

QY      62 NNNVEDYFKEGLEKTIIRAAKKARLEKTEAD 90
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      564 KGEI-----EAKRETLKEQREMDQLKSD 587

Search completed: June 21, 2005, 10:11:57
Job time : 11.9 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-11

Perfect score: 485

Sequence: 1 LKEIDSESDYVKEGFRAP.....KKAELEKTEADLKAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID    | Description        |
|------------|-------|---------------|--------|----------|--------------------|
| 1          | 485   | 100.0         | 417    | 2 Q9LAY3 | Q9lay3 streptococc |
| 2          | 459   | 96.7          | 619    | 2 Q54972 | Q54972 streptococc |
| 3          | 469   | 96.7          | 619    | 2 Q8DR10 | Q8dr10 streptococc |
| 4          | 447   | 92.2          | 415    | 2 Q9LAY1 | Q9lay1 streptococc |
| 5          | 414   | 85.4          | 739    | 2 Q9ROT4 | Q9rot4 streptococc |
| 6          | 414   | 85.4          | 820    | 2 Q9RQT1 | Q9rdt1 streptococc |
| 7          | 414   | 85.4          | 929    | 2 Q9KK19 | Q9kk19 streptococc |
| 8          | 414   | 85.4          | 929    | 2 Q9ZAY5 | Q9zay5 streptococc |
| 9          | 405   | 83.5          | 99     | 2 Q8KQK4 | Q8kk4 streptococc  |
| 10         | 398   | 82.1          | 224    | 2 Q9L575 | Q9l575 streptococc |
| 11         | 398   | 82.1          | 224    | 2 Q8GNS8 | Q8gns8 streptococc |
| 12         | 388   | 82.1          | 437    | 2 Q9LAY4 | Q9lay4 streptococc |
| 13         | 393   | 81.0          | 395    | 2 Q9LAY2 | Q9lay2 streptococc |
| 14         | 393   | 81.0          | 408    | 2 Q9LAY0 | Q9lay0 streptococc |
| 15         | 392   | 80.8          | 426    | 2 Q9LAY5 | Q9lay5 streptococc |
| 16         | 379.5 | 78.2          | 869    | 2 Q9KK27 | Q9kk27 streptococc |
| 17         | 363.5 | 74.9          | 225    | 2 Q9L591 | Q9l591 streptococc |
| 18         | 357.5 | 73.7          | 222    | 2 Q9L577 | Q9l577 streptococc |
| 19         | 357.5 | 73.7          | 262    | 2 Q9L576 | Q9l576 streptococc |
| 20         | 357.5 | 73.7          | 415    | 2 Q9LAY7 | Q9lay7 streptococc |
| 21         | 354.5 | 73.1          | 246    | 2 Q9L578 | Q9l578 streptococc |
| 22         | 350.5 | 72.3          | 416    | 2 Q9LAY8 | Q9lay8 streptococc |
| 23         | 349.5 | 72.1          | 255    | 2 Q9L581 | Q9l581 streptococc |
| 24         | 349.5 | 72.1          | 255    | 2 Q9L586 | Q9l586 streptococc |
| 25         | 348.5 | 71.9          | 406    | 2 Q9LAZ0 | Q9laz0 streptococc |
| 26         | 343.5 | 70.8          | 393    | 2 Q9LAZ3 | Q9laz3 streptococc |
| 27         | 342.5 | 70.6          | 394    | 2 Q9LAY6 | Q9lay6 streptococc |
| 28         | 342.5 | 70.6          | 395    | 2 Q9LAZ1 | Q9laz1 streptococc |
| 29         | 337.5 | 69.6          | 340    | 2 Q8KQK5 | Q8kk5 streptococc  |
| 30         | 334.5 | 69.0          | 207    | 2 Q8GNS9 | Q8gns9 streptococc |
| 31         | 332.5 | 68.6          | 237    | 2 Q9L592 | Q9l592 streptococc |

|    |       |      |     |          |                    |
|----|-------|------|-----|----------|--------------------|
| 32 | 332.5 | 68.6 | 395 | 2 Q9LAY9 | Q9lay9 streptococc |
| 33 | 326.5 | 67.3 | 194 | 2 Q9LSB5 | Q9lsb5 streptococc |
| 34 | 326.5 | 67.3 | 218 | 2 Q6UEB2 | Q6ueb2 streptococc |
| 35 | 326.5 | 67.3 | 233 | 2 Q9L568 | Q9l568 streptococc |
| 36 | 326.5 | 67.3 | 236 | 2 Q9L569 | Q9l569 streptococc |
| 37 | 326.5 | 67.3 | 243 | 2 Q9L564 | Q9l564 streptococc |
| 38 | 326.5 | 67.3 | 243 | 2 Q9L567 | Q9l567 streptococc |
| 39 | 326.5 | 67.3 | 244 | 2 Q9L565 | Q9l565 streptococc |
| 40 | 326.5 | 67.3 | 247 | 2 Q9L566 | Q9l566 streptococc |
| 41 | 326.5 | 67.3 | 249 | 2 Q9L570 | Q9l570 streptococc |
| 42 | 326.5 | 67.3 | 254 | 2 Q9L563 | Q9l563 streptococc |
| 43 | 326.5 | 67.3 | 401 | 2 Q9LAZ2 | Q9laz2 streptococc |
| 44 | 224.5 | 46.3 | 653 | 2 Q34097 | Q34097 streptococc |
| 45 | 201   | 41.4 | 107 | 2 Q8KQK2 | Q8kk2 streptococc  |

ALIGNMENTS

RESULT 1

|        |   |              |      |         |  |
|--------|---|--------------|------|---------|--|
| Q9LAY3 |   |              |      |         |  |
| ID     | Q9LAY3  | PRELIMINARY; | PRT; | 417 AA. |  |
| AC     | Q9LAY3;   |              |      |         |  |
| DT     | 01-OCT-2000 (Tremblrel. 15, Created)  |              |      |         |  |
| DT     | 01-OCT-2000 (Tremblrel. 15, Last sequence update)   |              |      |         |  |
| DT     | 01-OCT-2003 (Tremblrel. 25, Last annotation update)   |              |      |         |  |
| DE     | PepA (Fragment).  |              |      |         |  |
| GN     | Name=pspA;  |              |      |         |  |
| OS     | Streptococcus pneumoniae.   |              |      |         |  |
| OC     | Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  |              |      |         |  |
| OC     | Streptococcus.  |              |      |         |  |
| OX     | NCBI_TaxID=1313;  |              |      |         |  |
| RN     | [1]   |              |      |         |  |
| RP     | SEQUENCE FROM N.A.  |              |      |         |  |
| RC     | STRAIN=EF10197;   |              |      |         |  |
| RX     | MEDLINE=20448953; PubMed=10992499;  |              |      |         |  |
| RX     | DOI=10.1128/IAI.68.10.5889-5900.2000;   |              |      |         |  |
| RA     | Hollingshead S.K., Becker R., Briles D.B.;  |              |      |         |  |
| RT     | "Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae."; |              |      |         |  |
| RL     | Infect. Immun. 68:5889-5900(2000).  |              |      |         |  |
| DR     | EMBL; AF071812; AAF27708.1; -   |              |      |         |  |
| DR     | HSSP; P00192; 256B. 417   |              |      |         |  |
| FT     | NON_TER 417   |              |      |         |  |
| SQ     | SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;  |              |      |         |  |

Query Match 100.0%; Score 485; DB 2; Length 417;  
Best Local Similarity 100.0%; Pred. No. 6,5e-23;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |
|----|-----|---|
| QY | 1   | LKEIDSESDYVKEGFRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEDLQKAAE 60  |
| Db | 213 | LKEIDSESDYVKEGFRAPLQSELDKQAKLSKLEELSDKIDELDAETAKLEDLQKAAE 272 |
| QY | 61  | ENNVEDYFKEGKTIATAKKALEKTEADLKAVNEPE 99                        |
| Db | 273 | ENNVEDYFKEGKTIATAKKALEKTEADLKAVNEPE 311                       |

RESULT 2

|        |  |              |      |         |  |
|--------|--|--------------|------|---------|--|
| Q54972 |  |              |      |         |  |
| ID     | Q54972   | PRELIMINARY; | PRT; | 619 AA. |  |
| AC     | Q54972;  |              |      |         |  |
| DT     | 01-NOV-1996 (Tremblrel. 01, Created)                     |              |      |         |  |
| DT     | 01-NOV-1996 (Tremblrel. 01, Last sequence update)        |              |      |         |  |
| DT     | 01-MAR-2004 (Tremblrel. 26, Last annotation update)      |              |      |         |  |
| DE     | Pneumococcal surface protein A precursor.                |              |      |         |  |
| GN     | Name=pspA;   |              |      |         |  |
| OS     | Streptococcus pneumoniae.                                |              |      |         |  |
| OC     | Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; |              |      |         |  |
| OC     | Streptococcus.   |              |      |         |  |
| OX     | NCBI_TaxID=1313;   |              |      |         |  |
| RN     | [1]  |              |      |         |  |

```
RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis.";
RL J. Bacteriol. 174:601-609(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Signal.
FT SIGNAL 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 9.3e-22;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDAKQAKLSKLELSKDIDELDAETAKLEDQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKDIDELDAETAKLEDQKAAE 282

QY 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKAVNEPE 99
Db 283 ENNVEDYFKGLEKTTAAKAELEKTEADLKAVNEPE 321

RESULT 3
Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Surface protein pspA.
GN Name=pspA; OrderedLocNames=spr0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Albom W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glaes J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicass T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
```

```
KW Complete proteome.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 96.7%; Score 469; DB 2; Length 619;
Best Local Similarity 96.0%; Pred. No. 9.3e-22;
Matches 95; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDAKQAKLSKLELSKDIDELDAETAKLEDQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKDIDELDAETAKLEDQKAAE 282

QY 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKAVNEPE 99
Db 283 ENNVEDYFKGLEKTTAAKAELEKTEADLKAVNEPE 321

RESULT 4
Q9LAY1 PRELIMINARY; PRT; 415 AA.
AC Q9LAY1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WU2;
RX MEDLINE=204448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071814; AAF27710.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
DR NON_TER 415 415
SQ SEQUENCE 415 AA; 46075 MW; 213C1AF7FF21642F CRC64;

Query Match 92.2%; Score 447; DB 2; Length 415;
Best Local Similarity 91.9%; Pred. No. 1.6e-20;
Matches 91; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYVKEGFRAPLQSELDAKQAKLSKLELSKDIDELDAETAKLEDQKAAE 60
Db 221 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKDIDELDAETAKLEDQKAVE 280

QY 61 ENNVEDYFKGLEKTTAAKAELEKTEADLKAVNEPE 99
Db 281 ENNVEDYFKGLEKTTAAKAELEKTEADLKAVNEPE 319

RESULT 5
Q9RQT4 PRELIMINARY; PRT; 739 AA.
AC Q9RQT4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068647; AAF13457.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EE2F2F676ABF989 CRC64;

Query Match 85.4%; Score 414; DB 2; Length 739;
Best Local Similarity 86.9%; Pred. No. 3e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60
DB 537 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 596

QY 61 ENNVVDYFKEGLEKTIKAAKAELEKTEADLKKAVNEPE 99
DB 597 GNNVVEAYFKEGLEKTIKAAKAELEKTEADLKKAVNEPE 635

RESULT 6
Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; 1IC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 85.4%; Score 414; DB 2; Length 820;
Best Local Similarity 86.9%; Pred. No. 3.3e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60
DB 530 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 589

QY 61 ENNVVDYFKEGLEKTIKAAKAELEKTEADLKKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTIKAAKAELEKTEADLKKAVNEPE 628

RESULT 7
Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=brf10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002)
DR EMBL; AF154037; AAF73809.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 85.4%; Score 414; DB 2; Length 929;
Best Local Similarity 86.9%; Pred. No. 3.7e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 60
DB 530 LKEIDSESDYVKEGFRAPLOSELDAKQAKLSKLEELSDKIDELDAETAKLEDQKAAE 589

QY 61 ENNVVDYFKEGLEKTIKAAKAELEKTEADLKKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTIKAAKAELEKTEADLKKAVNEPE 628

RESULT 8
Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooke-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP65; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; YSRK signal; 1.
DR TIGRFAMs; TIGR01168; YSRK_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match      85.4%; Score 414; DB 2; Length 929;
Best Local Similarity 86.9%; Pred. No. 3.7e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEDLKAAE 60
    |||||
Db 530 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEDLKAAE 589
    |||||

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99
    |||||
Db 590 GNNVVEAYFKEGLEKTTAAKKAELKTEADLKAVNEPE 628
    |||||

RESULT 9
Q8KQK4
ID Q8KQK4 PRELIMINARY; PRT; 99 AA.
AC Q8KQK4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082388; AAL92493.1; -.
FT NON_TER 1
FT NON_TER 99
FT NON_TER 99
SQ SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match      83.5%; Score 405; DB 2; Length 99;
Best Local Similarity 85.9%; Pred. No. 1.8e-18;
Matches 85; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEDLKAAE 60
    |||||
Db 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAEIAEVLQKDAE 60
    |||||

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99
    |||||
Db 61 GNNVVEAYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99
    |||||

```

```

RESULT 10
Q9L575
ID Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packiam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match      83.5%; Score 405; DB 2; Length 249;
Best Local Similarity 84.8%; Pred. No. 4.2e-18;
Matches 84; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAEIAKLEDLKAAE 60
    |||||
Db 74 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSKDIDELDAEIAEVLQKDAE 133
    |||||

Qy 61 ENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 99
    |||||
Db 134 GNNVVEAYFKEGLEKTTAAKKAELKTEADLKAVNEPE 172
    |||||

RESULT 11
Q8GNS8
ID Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "genotypes of invasive pneumococcal isolates recently recovered from
RT Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
DR HSP; P00192; 1APC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 224
FT NON_TER 224

```



```
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 82.1%; Score 398; DB 2; Length 224;
Best Local Similarity 83.8%; Pred. No. 1e-17;
Matches 83; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 17 LKIDNESDSYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 76
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENNVNEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 77 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 115

RESULT 12
Q9LAY4
ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PpA (Fragment).
GN Name=ppaA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC STRAIN=EL34;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PpA: mosaic genes and evidence for past recombination
   in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW binding_1; 1.
FT NON TER 437 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 82.1%; Score 398; DB 2; Length 437;
Best Local Similarity 84.8%; Pred. No. 1.9e-17;
Matches 84; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKHVVLQKAAE 294
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENNVNEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 295 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 333

RESULT 13
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE PpA (Fragment).
GN Name=ppaA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC STRAIN=EL34;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PpA: mosaic genes and evidence for past recombination
   in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW binding_1; 1.
FT NON TER 437 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 82.1%; Score 398; DB 2; Length 437;
Best Local Similarity 84.8%; Pred. No. 1.9e-17;
Matches 84; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKHVVLQKAAE 294
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENNVNEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 295 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 333

RESULT 14
Q9LAY0
ID Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE PpA (Fragment).
GN Name=ppaA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PpA: mosaic genes and evidence for past recombination
   in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON TER 408 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 81.0%; Score 393; DB 2; Length 408;
Best Local Similarity 82.8%; Pred. No. 3.7e-17;
Matches 82; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 228 LKEIDSDSDYKGFAPLOSLEDAKQAKLSKLELSKDIDELDAETAKLEDOLKAAE 287
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ENNVNEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 288 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKAVDEPE 326

RESULT 15
Q9LAY5
ID Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
```

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PSpA (Fragment).
GN Name=pspa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DEL5;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.168.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR HSSP; P00192; IMGT.
DR InterPro; IPR011047; Quin_abc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CB6634 CRC64;

Query Match      80.8%; Score 392; DB 2; Length 426;
Best Local Similarity 82.8%; Pred. No. 4.4e-17;
Matches 82; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 1 LKEIDSDSEDYVVEGFRAPLQSELDKQAKLSKLELSKIDELDAEIAKLEDLKAAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 215 LKDINESDSEDYVVEGLRAPLQSELDYKAKLLKLELSKIEELDAEIAELEVLQKDAE 274
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 ENNVVEDYFKEGLEKTTAAKKAELKTEADLKKAVNEPE 99
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 275 GNNVVEAYFKEGLEKTTAEKKAELKAEADLUKKAVDPE 313
```

Search completed: June 21, 2005, 10:22:10  
Job time : 63.3194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 74.5918 Seconds  
(without alignments)

518.502 Million cell updates/sec

Title: US-10-674-755-12

Perfect score: 489

Sequence: 1 LKEDISESDYAKGFRAP.....KKALEKTEADLKAVNEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 469.5 | 96.0        | 198    | 7     | ABW02615 Rx1c pneu  |
| 2          | 469.5 | 96.0        | 315    | 2     | AAY04375 Streptoco  |
| 3          | 469.5 | 96.0        | 619    | 2     | AAR63437 Pneumoco   |
| 4          | 469.5 | 96.0        | 619    | 2     | AAR87598 Pneumoco   |
| 5          | 469.5 | 96.0        | 619    | 2     | AAR86911 Pneumoco   |
| 6          | 469.5 | 96.0        | 619    | 2     | AAY41838 Streptoco  |
| 7          | 469.5 | 96.0        | 619    | 5     | AAE18782 S. pneumo  |
| 8          | 469.5 | 96.0        | 619    | 6     | ABU45778 Protein e  |
| 9          | 469.5 | 96.0        | 619    | 8     | ADO52126 Streptoco  |
| 10         | 469.5 | 96.0        | 648    | 2     | AAW70336 Pneumoco   |
| 11         | 469.5 | 96.0        | 648    | 2     | AAW62274 Streptoco  |
| 12         | 469.5 | 96.0        | 648    | 2     | AAY41837 Streptoco  |
| 13         | 469.5 | 96.0        | 648    | 2     | AAW87879 A pneumoc  |
| 14         | 469.5 | 96.0        | 653    | 2     | AAW92456 S. pneumo  |
| 15         | 469.5 | 96.0        | 684    | 2     | AAAR73912 Streptoco |
| 16         | 469.5 | 96.0        | 8991   | 6     | ABU08487 S. pneumo  |
| 17         | 466.5 | 95.4        | 198    | 2     | AAW14581 Streptoco  |
| 18         | 462.5 | 94.6        | 204    | 2     | AAW14571 Streptoco  |
| 19         | 462.5 | 94.6        | 204    | 7     | ABW02605 Efi1019c p |
| 20         | 449.5 | 91.9        | 653    | 2     | AAAR27150 PspA frag |
| 21         | 447   | 91.4        | 289    | 2     | AAW62276 Streptoco  |
| 22         | 447   | 91.4        | 289    | 2     | AAW41840 Streptoco  |
| 23         | 447   | 91.4        | 289    | 2     | AAW87910 Protein s  |
| 24         | 447   | 91.4        | 289    | 2     | AAW92458 S. pneumo  |
| 25         | 442.5 | 90.5        | 195    | 2     | AAW14591 Streptoco  |

|    |       |      |      |   |          |                    |
|----|-------|------|------|---|----------|--------------------|
| 26 | 442.5 | 90.5 | 195  | 7 | ABW02625 | Abw02625 Wu2c pneu |
| 27 | 426.5 | 87.2 | 1231 | 6 | ABU08490 | Abu08490 Fragment  |
| 28 | 423.5 | 86.6 | 623  | 6 | ABU08494 | Abu08494 Fragment  |
| 29 | 402.5 | 82.3 | 170  | 7 | ABW02614 | Abw02614 Rct135c p |
| 30 | 402.5 | 82.3 | 181  | 7 | ABW02596 | Abw02596 0922134c  |
| 31 | 402.5 | 82.3 | 865  | 6 | ABU08489 | Abu08489 S. pneumo |
| 32 | 402.5 | 82.3 | 929  | 2 | AAW14593 | AAW14593 Streptoco |
| 33 | 402.5 | 82.3 | 929  | 2 | AAW14593 | AAW14593 Streptoco |
| 34 | 399.5 | 81.7 | 188  | 7 | AAW14580 | AAW14580 Streptoco |
| 35 | 399.5 | 81.7 | 188  | 7 | AAW14580 | AAW14580 Streptoco |
| 36 | 392.5 | 80.3 | 204  | 2 | AAW14578 | AAW14578 Streptoco |
| 37 | 392.5 | 80.3 | 204  | 2 | AAW14578 | AAW14578 Streptoco |
| 38 | 390.5 | 79.9 | 588  | 6 | ABU08491 | Abu08491 Coiled co |
| 39 | 390.5 | 79.9 | 588  | 6 | ABU08491 | Abu08491 Coiled co |
| 40 | 388   | 79.3 | 180  | 2 | AAW14562 | AAW14562 Streptoco |
| 41 | 385   | 78.7 | 187  | 2 | AAW14579 | AAW14579 Streptoco |
| 42 | 375.5 | 76.8 | 206  | 2 | AAW14574 | AAW14574 Streptoco |
| 43 | 375.5 | 76.8 | 206  | 2 | ABW02608 | Abw02608 Db15c pne |
| 44 | 351   | 71.8 | 550  | 8 | ADK48356 | Adk48356 Streptoco |
| 45 | 351   | 71.8 | 550  | 8 | ADR95223 | Adr95223 Novel S.  |

#### ALIGNMENTS

##### RESULT 1

ABW02615

ID ABW02615 standard; protein; 198 AA.

XX AC ABW02615;

XX DT 12-FEB-2004 (first entry)

XX DE Rx1c pneumococcal surface protein A (PspA) central region.

XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

XX KW immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX PN US6592876-B1.

XX PD 15-JUL-2003.

XX PF 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for

XX detecting the presence of Streptococcus pneumoniae or its strain,

XX comprises at least two different full length isolated gene encoding

XX pneumococcal surface protein A.

XX Example 6; SEQ ID NO 61; 121pp; English.

XX The present invention relates to an immunological composition comprising

XX at least 2 different full length isolated genes encoding pneumococcal

XX surface protein A (PspAs) from different groups based on restriction

XX fragment polymorphism analysis. The invention is useful for obtaining

XX expression products by recombinant techniques to detect, determine,

XX isolate or diagnose the presence of Streptococcus pneumoniae or its

XX strain. The expression product is useful for preparing antigenic,

XX immunological or vaccine compositions, for eliciting antibodies, an

XX immunological response (other than or additional to antibodies) or a

XX protective response (including antibody or other immunological response

XX by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Rx1c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention

XX SQ Sequence 198 AA;

Query Match 96.0%; Score 469.5; DB 7; Length 198;  
 Best Local Similarity 98.0%; Pred. No. 4.7e-34;  
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60  
 |||||  
 Db 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 59  
 |||||

Qy 61 EENNVEDYFKGLEKTIAAKKALEKTEADLKAVNEPE 100  
 |||||

Db 60 EENNVEDYFKGLEKTIAAKKALEKTEADLKAVNEPE 99  
 |||||

#### RESULT 2

AA04375  
 ID AA04375 standard; protein; 315 AA.

XX AC AA04375;

DT 23-JUN-1999 (first entry)

XX Streptococcus pneumoniae PspA protein sequence.

XX Streptococcus pneumoniae; pspA; pneumococcal; surface protein; vaccine;  
 KW immunological; infection.

XX Streptococcus pneumoniae.

OS Synthetic.

XX WO9914333-A2.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US019740.

XX 18-SEP-1997; 97US-00932982.

XX (INNR ) PASTEUR MERIEUX CONNAUGHT.

XX Becker R, Gray M, Pyle D;

XX WPI; 1999-229537/19.

DR N-PSDB; AAX33124.

XX DNA encoding PspA molecule with modified internal translational  
 PT initiation sites.

XX Disclosure; Page; 36pp; English.

XX The present sequence represents a pneumococcal surface protein A (PspA)  
 CC molecule where internal naturally occurring translational initiation  
 CC sites have been modified or eliminated so that expression of the DNA  
 CC sequence results in a single form of PspA. The PspA nucleotide sequence  
 CC can be used to transform a unicellular host to produce the PspA protein.  
 CC The PspA protein can be used in an immunological composition for treating  
 CC or preventing S. pneumoniae infection especially in a child. Antibodies  
 CC to the PspA protein can also be used to treat S. pneumoniae infection.  
 CC The immunogenic peptides are designed to confer broad protection against  
 CC diverse pneumococcal strains. N.B. The present sequence is not given in  
 CC the specification but is encoded by the sequence given in AAX33124

XX SQ Sequence 315 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 315;  
 Best Local Similarity 98.0%; Pred. No. 8.2e-34;  
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60  
 |||||  
 Db 193 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 251  
 |||||

Qy 61 EENNVEDYFKGLEKTIAAKKALEKTEADLKAVNEPE 100  
 |||||

Db 252 EENNVEDYFKGLEKTIAAKKALEKTEADLKAVNEPE 291  
 |||||

#### RESULT 3

AA63437  
 ID AA63437 standard; protein; 619 AA.

XX AC AA63437;

DT 09-SEP-2004 (revised)

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 19-JUL-1995 (first entry)

XX Pneumococcal surface protein A from S.pneumoniae Rx1.

XX Pneumococcal surface protein A; pspA; Streptococcus; PCR; pneumococcal;  
 KW primer; protection-eliciting epitope; epitope; vaccine; amplify.

OS Streptococcus pneumoniae.

OS Unidentified.

XX Key Location/Qualifiers

FT Protein 192..260  
 FT /note= "protein fragment of Claim 1"

XX EP622081-A2.

XX 02-NOV-1994.

XX 19-APR-1994; 94EP-00302767.

XX 20-APR-1993; 93US-00048896.

XX (UABR-) UAB RES FOUND.

XX Briles DE, Yother JL, McDaniel LS;

XX WPI; 1994-359522/45.

XX N-PSDB; AAQ78131.

XX regions of Pneumococcal surface protein A - derived from the Rx1 PspA  
 PT strain, for the preparation of cross-reactive vaccines for the prevention  
 PT of pneumococcal infections.

XX Disclosure; Page 13-16; 26pp; English.

XX The amino acid sequence of the novel Pneumococcal surface protein A  
 CC (PspA) from Streptococcus pneumoniae strain Rx1. The gene was PCR  
 CC amplified from S.pneumoniae genomic DNA using the primers AAQ78132-5. The  
 CC gene was used to derive truncated peptide fragments containing protection  
 CC -eliciting epitopes for use in vaccines against pneumococcal diseases.  
 CC The epitopic fragments are derived from amino acids 192-260 and C-  
 CC terminally contain a further 25 a.a. residues at both the N- and C-  
 CC terminal regions of the peptide. The epitopic peptide fragments may be  
 CC derived from different strains of S.pneumoniae and are homologous to the  
 CC Rx1 strain epitope. (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 16-OCT-2003 to standardise OS field)

XX Revised record issued on 09-SEP-2004 : Correction to feature table key

XX SQ Sequence 619 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 619;  
 Best Local Similarity 98.0%; Pred. No. 1.9e-33;  
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```
QY 1 LKEIDSESEDYAKGFRAPLOSQKLDKAKOAKLSKLELSKIDELDAETAKLEDOLKKA 60
Db 223 LKEIDSESEDYAKGFRAPLOSQKLDKAKKAKLSKLELSKIDELDAETAKLEDOLKKA 281
QY 61 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 4
AAR87598
ID AAR87598 standard; protein; 619 AA.
XX
AC AAR87598;
XX
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-JUL-1996 (first entry)
XX
DE Pneumococcal surface protein (PspA).
XX
XX PspA; pneumococcal surface protein; truncated; immunoprotective;
XX soluble fragment; insertion-duplication mutagenesis.
XX
OS Streptococcus pneumoniae; strain Rxl.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein 32..619 /label= signal_peptide
FT Region 32..319 /label= mature_protein
FT /label= alpha-helical coiled-coil region
FT /note= "contains a seven-residue periodicity"
FT Region 320..401
FT Region 402..421 /note= "proline-rich region"
FT Region 422..441 /note= "repeat region"
FT Region 442..461 /note= "repeat region"
FT Region 462..481 /note= "repeat region"
FT Region 482..501 /note= "repeat region"
FT Region 502..521 /note= "repeat region"
FT Region 522..541 /note= "repeat region"
FT Region 542..561 /note= "repeat region"
FT Region 562..581 /note= "repeat region"
FT Region 582..619 /note= "repeat region"
FT /note= "hydrophobic region starts in last repeat region
FT is potential membrane-spanning region"
XX
XX US5476929-A.
XX
XX 19-DEC-1995.
XX
XX 03-JUN-1993; 93US-00072070.
XX
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX
XX (UABR-) UAB RES FOUND.
XX
XX McDaniel LS, Yother JL, Briles DE;
XX
XX WPI; 1996-049021/05.
DR N-PSDB; AAT08979.
```

```
XX New pneumococcal surface protein A fragments - comprise proline-rich
PT region and/or repeat region, used for detection of Streptococcus
PT pneumoniae.
XX
XX Claim 1; Col 15-20; 23pp; English.
XX
XX The present sequence is that of PspA (pneumococcal surface protein A)
CC encoded by AAT08979. Through the technique of insertion-duplication
CC mutagenesis of the pspA gene of the strain Rxl of Streptococcus
CC pneumoniae with plasmids contg. cloned fragments of the pspA structural
CC gene, it has been possible to produce soluble fragments of PspA that are
CC secreted by pneumococci. The method can be used to provide an
CC immunoprotective truncated PspA protein. Primers and probes based on the
CC present sequence are claimed, and are useful for the detection of (at
CC least 32) S. pneumoniae strains. (Updated on 25-MAR-2003 to correct PF
CC field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 619 AA;
Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. NO. 1.9e-33;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESEDYAKGFRAPLOSQKLDKAKOAKLSKLELSKIDELDAETAKLEDOLKKA 60
Db 223 LKEIDSESEDYAKGFRAPLOSQKLDKAKKAKLSKLELSKIDELDAETAKLEDOLKKA 281
QY 61 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 5
AAR86911
ID AAR86911 standard; protein; 619 AA.
XX
AC AAR86911;
XX
XX 16-OCT-2003 (revised)
DT 11-MAY-1996 (first entry)
DE Pneumococcal surface protein A.
XX
XX Pneumococcal surface protein A; PspA; cross-protection; vaccine;
KW Streptococcus pneumoniae; probe; primer; polymerase chain reaction;
KW otitis media; meningitis; bacteraemia; pneumonia; epitope.
XX
OS Streptococcus pneumoniae; strain Rxl.
XX
FH Key Location/Qualifiers
FT Peptide 1..31 /label= sig_peptide
FT Region 32..288
FT /note= "N-terminal region is highly charged and includes
FT an alpha-helix structure"
FT 289..619
FT Region /note= "C-terminal region includes a proline-rich region
FT and a repeat region"
XX
XX AU9520112-A.
XX
XX 30-NOV-1995.
XX
XX 18-MAY-1995; 95AU-00020112.
XX
XX 20-MAY-1994; 94US-00246636.
XX 07-OCT-1994; 94US-00319795.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Yother JL, McDaniel LS;
XX
XX
```

```

DR WPI; 1996-030801/04.
DR N-PSDB; AAT07178.
XX Pneumococcal DNA primers and probes - amplify and detect cross-protective
PT epitope(s) from Streptococcus pneumoniae surface protein A.
XX Disclosure; Page 41-43; 61pp; English.
XX Surface protein A, PspA (AAR86911), of Streptococcus pneumoniae Rx1 is
CC the product of the pspA gene (AAT07178). PspA includes regions
CC comprising e.g. amino acids 182-588, 293-588 and 192-299, that elicit
CC cross-protection against challenge by multiple wild-type strains of S.
CC pneumoniae. These cross-reactive epitopes can be prepd. by expression of
CC DNA obt'd. by PCR amplification (see AAT07179-96), for use in vaccine
CC compns. (Updated on 16-OCT-2003 to standardise OS field)
XX Sequence 619 AA;
SQ
Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-33;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
Oy 1 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 281
Oy 61 EENNVEDYFKGLEKTIIAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGLEKTIIAAKAELEKTEADLKAVNEPE 321
RESULT 6
AAAY41838
ID AAAY41838 standard; protein; 619 AA.
XX
AC AAAY41838;
XX
DT 08-DEC-1999 (first entry)
XX
DE Streptococcus pneumoniae Rx1 PspA protein sequence.
XX
KW Streptococcus pneumoniae Rx1; PspA; immunoprotective; vaccine; diagnosis;
KW infection; pneumococcal surface protein A.
XX
OS Streptococcus pneumoniae.
XX
PN US965400-A.
XX
PD 12-OCT-1999.
XX
PF 23-MAY-1994; 9AUS-00247491.
XX
PR 15-FEB-1991; 9IUS-00656773.
XX
PR 12-FEB-1992; 9ZUS-00835698.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Yother JL, Briles DE;
XX
XX WPI; 1999-579913/49.
DR N-PSDB; AAZ25063.
XX
XX DNA encoding a truncated pneumococcal surface protein A used in the
PT development of pneumococcal infections.
XX
PS Claim 1; Fig 3; 27pp; English.
XX
XX The present sequence represents Streptococcus pneumoniae Rx1
CC immunoprotective Pneumococcal surface protein A (PspA). The present
CC invention also describes a method of forming the immunoprotective
CC truncated PspA, comprising incorporating a vector comprising the isolated
CC DNA molecule encoding PspA (I), into a bacterium via transformation. (I)
CC is used to design primers which are capable of detecting a large number

```

---

```

CC of S. pneumoniae strains, which in turn can be used to diagnose
CC pneumococcal infection in mammals (e.g. humans), independent of the
CC strain which has caused it. The PspA protein is used to develop a vaccine
CC against pneumococcal infection comprising, as an immunologically-active
CC component, a live attenuated or killed bacteria containing a gene coding
CC for the truncated form of PspA
XX
XX Sequence 619 AA;
SQ
Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-33;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
Oy 1 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAA 281
Oy 61 EENNVEDYFKGLEKTIIAAKAELEKTEADLKAVNEPE 100
Db 282 EENNVEDYFKGLEKTIIAAKAELEKTEADLKAVNEPE 321
RESULT 7
AAE18782
ID AAE18782 standard; protein; 619 AA.
XX
AC AAE18782;
XX
DT 17-MAY-2002 (first entry)
XX
DE Streptococcus pneumoniae Rx1 strain pneumococcal surface protein A (PspA).
XX
KW Coiled-coil structural scaffold; heptad repeat; epitope; immune response;
KW cell-mediated immunity; microbial infection; cross-protection; therapy;
KW antimicrobial; vaccine; pneumococcal surface protein A; PspA.
XX
OS Streptococcus pneumoniae.
XX
XX Key Location/Qualifiers
FH Domain 1..314
FT /label= Helical_domain
FT Region 1..303
FT /note= "N-terminal region"
FT Region 38..44
FT /note= "Immunogenic region 3"
FT Region 40..46
FT /note= "Immunogenic region 5"
FT Region 75..80
FT /note= "Immunogenic region 29"
FT Region 82..87
FT /note= "Immunogenic region 52"
FT Region 96..101
FT /note= "Immunogenic region 66"
FT Region 114..119
FT /note= "Immunogenic region 73"
FT Region 130..135
FT /note= "Immunogenic region 78"
FT Region 137..142
FT /note= "Immunogenic region 89"
FT Region 140..145
FT /note= "Immunogenic region 91"
FT Region 152..156
FT /note= "Immunogenic region 95"
FT Domain 153..170
FT /label= Coiled_coil_motif
FT Region 161..164
FT /note= "Immunogenic region 101"
FT Region 166..170
FT /note= "Immunogenic region 116"
FT Region 173..177
FT /note= "Immunogenic region 122"
FT Region 176..180
FT /note= "Immunogenic region 123"

```



CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 619 AA;

Query Match 96.0%; Score 469.5; DB 6; Length 619;  
 Best Local Similarity 98.0%; Pred. No. 1.9e-33;  
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSDKIDELDAEIAKLEDQKAA 60  
 |||||  
 DB 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSDKIDELDAEIAKLEDQK-CAA 281  
 |||||

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100  
 |||||  
 DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321  
 |||||

RESULT 9  
 ADO52126  
 ID ADO52126 standard; protein; 619 AA.

XX AC ADO52126;

XX DT 12-AUG-2004 (first entry)

XX STREPTOCOCCUS PNEUMONIAE Rx1 PspA protein.

XX IMMUNOGENIC COMPOSITION; vaccine; Th2-type immune response;  
 KW pneumococcal surface protein A; PspA.

XX STREPTOCOCCUS PNEUMONIAE.

Key Location/Qualifiers  
 FH Peptide 1..31  
 FT /label= Signal\_peptide  
 FT Protein 32..619  
 FT /note= "S. pneumoniae Rx1 mature PspA protein"

XX US20040101531-A1.

XX 27-MAY-2004.

XX 15-APR-2003; 2003US-00414532.

XX 16-APR-2002; 2002US-0372710P.

XX (CURTIS) CURTISS R.

XX (KANG) KANG H Y.

XX CURTISS R, KANG HY;

XX WPI; 2004-399655/37.

XX N-PSDB; ADO52125.

XX New vaccine comprising a live attenuated strain of pathogenic gram-  
 PT negative bacteria, useful in eliciting a Th2-type immune response in a  
 PT vertebrate against pathogens, e.g., helminths, fungi, viruses, protozoans  
 PT or bacteria.

XX Example 8; SEQ ID NO 72; 94pp; English.

XX The invention relates to immunogenic compositions and vaccines comprising  
 CC a live attenuated strain of pathogenic gram negative bacteria that  
 CC secretes an antigen. The vaccine is useful in eliciting a Th2-type immune  
 CC response in a vertebrate against pathogens, e.g., helminths, fungi,  
 CC viruses, protozoans or bacteria. The present sequence is Streptococcus  
 CC pneumoniae Rx1 pneumococcal surface protein A (PspA). This sequence is  
 CC used in the exemplification of the invention.

XX  
 SQ Sequence 619 AA;

Query Match 96.0%; Score 469.5; DB 8; Length 619;  
 Best Local Similarity 98.0%; Pred. No. 1.9e-33;  
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSDKIDELDAEIAKLEDQKAA 60  
 |||||  
 DB 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSDKIDELDAEIAKLEDQK-CAA 281  
 |||||

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 100  
 |||||

DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKKAVNEPE 321  
 |||||

RESULT 10

AAW70336

ID AAW70336 standard; protein; 648 AA.

XX AC AAW70336;

XX DT 18-NOV-1998 (first entry)

XX PNEUMOCOCCAL SURFACE PROTEIN A (PSPA).

XX PNEUMOCOCCAL SURFACE PROTEIN A gene; PspA; PspA epitope; vaccine;  
 KW insertion-duplication mutagenesis; otitis media; meningitis; bacteraemia;  
 KW pneumonia.

XX STREPTOCOCCUS PNEUMONIAE.

Key Location/Qualifiers  
 FH Peptide 1..31  
 FT /note= "Signal peptide"  
 FT Protein 32..648  
 FT /note= "PspA"  
 FT Region 32..319  
 FT /note= "alpha-helical coil region representing the  
 FT truncated PspA of the invention"  
 FT Misc-difference 647  
 FT /note= "Encoded by AGG"

XX US5804193-A.

XX 08-SEP-1998.

XX 17-MAR-1994; 94US-00214222.

XX 15-FEB-1991; 91US-00656773.

XX 12-FEB-1992; 92US-00835698.

XX (UABR-) UAB RES FOUND.

XX BRILES DE, YOTHER JL;

XX WPI; 1998-505588/43.

XX N-PSDB; AAV33264.

XX Truncated pneumococcal surface protein - useful in vaccines against  
 PT pneumococcal infection.

XX Example 3; Fig 3A-3C; 22pp; English.

XX The present sequence represents the Streptococcus pneumoniae Rx1  
 CC pneumococcal surface protein A (PSPA). The invention provides a purified  
 CC truncated form of PspA, formed by an insertion-duplication mutagenesis  
 CC technique, comprising of the first 288 N-terminal residues of the mature  
 CC form of wild-type PspA (AAW70336). The truncated PspA contains  
 CC immunoprotective epitopes of PspA. The invention claims for a vaccine  
 CC against pneumococcal infection, comprising live-attenuated or killed *S.*  
 CC pneumoniae, containing the gene coding for the truncated PspA protein.  
 CC The truncated protein, optionally conjugated to a poorly immunogenic or



CC nonimmunogenic molecule, is claimed to be useful in vaccines against  
 CC pneumococcal infection, especially otitis media, meningitis, bacteraemia  
 CC and pneumonia  
 XX  
 SQ Sequence 648 AA;  
 Query Match 96.0%; Score 469.5; DB 2; Length 648;  
 Best Local Similarity 98.0%; Pred. No. 2e-33;  
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLELSKDIDELDAETAKLEDQLKAA 60  
 |||||  
 Db 223 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLELSKDIDELDAETAKLEDQ-L-KAA 281  
 |||||  
 QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 100  
 |||||  
 Db 282 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 321  
 |||||  
 RESULT 11  
 AAW62274  
 ID AAW62274 standard; protein; 648 AA.  
 XX  
 AC AAW62274;  
 XX  
 DT 22-SEP-1998 (first entry)  
 XX  
 DE Streptococcus pneumoniae pspA protein.  
 XX  
 DE Streptococcus pneumoniae strain Rxi; pspA; immunoprotective; immunogen;  
 KW Streptococcal surface protein A; cholera toxin B subunit; fusion protein;  
 KW antigenic.  
 KW  
 XX Streptococcus pneumoniae.  
 OS  
 FT Key Location/Qualifiers  
 FT Misc-difference 619. .620  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 621. .622  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 625. .626  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 630. .631  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 632. .633  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT  
 XX US5965400-A.  
 PN  
 XX 12-OCT-1999.  
 XX  
 PD 23-MAY-1994; 94US-00247491.  
 XX  
 PF 15-FEB-1991; 91US-00656773.  
 XX  
 PR 12-FEB-1992; 92US-00835698.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA  
 Yother JL, Briles DE;  
 WPI; 1998-311399/27.  
 DR N-PSDB; AAV39470.  
 DR  
 PT Truncated pneumococcal surface protein and cholera toxin B sub-unit  
 PT fusion protein - useful as an immunogen against Streptococcus pneumoniae.  
 PS  
 PS Claim 1; Fig 3; 22pp; English.  
 XX  
 CC The present sequence represents the pneumococcal surface protein A (PspA)  
 CC protein from Streptococcus pneumoniae. A recombinant DNA molecule has  
 CC been developed which encodes a fusion protein comprising a truncated form  
 CC of PspA and cholera toxin B subunit (CTB), where the DNA molecule  
 CC comprises a nucleotide sequence encoding the truncated PspA linked by an  
 CC in-frame genetic fusion to a ctaB gene, and where the truncated PspA  
 CC contains immunoprotective epitopes and up to 90% of the whole PspA  
 CC protein, except for the cell membrane anchor region. The fusion protein  
 CC is useful for providing an immunogen to protect neonates and children  
 CC against S.pneumoniae. Most antigenic proteins of this strain are not  
 CC immunogenic enough to provide protection. The antigenic epitopes of the  
 CC fusion protein are directed against capsular polysaccharide antigens of  
 CC S.pneumoniae, specifically it contains the protective epitopes of PspA.  
 CC The protein can also be used in solid-phase immunosorbent assays, since  
 CC it is readily bound to supports coated with monosialoganglioside GM1. The  
 CC fusion protein is more immunogenic against S.pneumoniae than using PspA

CC alone as the immunogen  
 XX  
 SQ Sequence 648 AA;  
 Query Match 96.0%; Score 469.5; DB 2; Length 648;  
 Best Local Similarity 98.0%; Pred. No. 2e-33;  
 Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLELSKDIDELDAETAKLEDQLKAA 60  
 |||||  
 Db 223 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLELSKDIDELDAETAKLEDQ-L-KAA 281  
 |||||  
 QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 100  
 |||||  
 Db 282 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 321  
 |||||  
 RESULT 12  
 AAY41837  
 ID AAY41837 standard; protein; 648 AA.  
 XX  
 AC AAY41837;  
 XX  
 DT 08-DEC-1999 (first entry)  
 XX  
 DE Streptococcus pneumoniae Rxi PspA protein sequence.  
 XX  
 DE Streptococcus pneumoniae Rxi; PspA; immunoprotective; vaccine; diagnosis;  
 KW infection; pneumococcal surface protein A.  
 KW  
 XX Streptococcus pneumoniae.  
 OS  
 FT Key Location/Qualifiers  
 FT Misc-difference 619. .620  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 621. .622  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 625. .626  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 630. .631  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT Misc-difference 632. .633  
 FT /note= "a stop codon is present in the nucleotide  
 sequence at this position"  
 FT  
 XX US5965400-A.  
 PN  
 XX 12-OCT-1999.  
 XX  
 PD 23-MAY-1994; 94US-00247491.  
 XX  
 PF 15-FEB-1991; 91US-00656773.  
 XX  
 PR 12-FEB-1992; 92US-00835698.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA  
 Yother JL, Briles DE;  
 WPI; 1999-579913/49.  
 DR N-PSDB; AA225063.  
 DR  
 PT DNA encoding a truncated pneumococcal surface protein A used in the  
 PT development of pneumococcal infections.  
 PS  
 PS Claim 1; Fig 3; 27pp; English.  
 XX  
 CC The present sequence represents Streptococcus pneumoniae Rxi  
 CC immunoprotective Pneumococcal surface protein A (PspA). The present  
 CC invention also describes a method of forming the immunoprotective

CC truncated PspA, comprising incorporating a vector comprising the isolated  
 CC DNA molecule encoding PspA (I), into a bacterium via transformation. (I)  
 CC is used to design primers which are capable of detecting a large number  
 CC of *S. pneumoniae* strains, which in turn can be used to diagnose  
 CC pneumococcal infection in mammals (e.g. humans), independent of the  
 CC strain which has caused it. The PspA protein is used to develop a vaccine  
 CC against pneumococcal infection comprising, as an immunologically-active  
 CC component, a live attenuated or killed bacteria containing a gene coding  
 CC for the truncated form of PspA  
 XX  
 SQ Sequence 648 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 648;  
 Best Local Similarity 98.0%; Pred. No. 2e-33;  
 Matches 98; Conservative 1; Mismatches 1; Gaps 1;  
 QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 60  
 DB 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 281  
 QY 61 EENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEP 100  
 DB 282 EENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEP 321

RESULT 13  
 AAW8789  
 ID AAW8789 standard; protein; 648 AA.

AC AAW8789;  
 XX  
 XX 19-MAR-1999 (first entry)  
 XX  
 XX A pneumococcal surface protein (PspA).  
 XX Pneumococcal surface protein; PspA; Streptococcus pneumoniae Rx1;  
 KW truncated protein; epitope-containing region; vaccine;  
 KW pneumococcal infection.  
 XX Streptococcus pneumoniae.  
 XX US5856170-A.  
 XX  
 XX 05-JAN-1999.  
 XX  
 XX 06-JUN-1995; 95US-00467852.  
 XX  
 XX 15-FEB-1991; 91US-00656773.  
 XX 12-FEB-1992; 92US-00835698.  
 XX 23-MAY-1994; 94US-00247491.  
 XX (UABR-) UAB RES FOUND.  
 XX  
 XX Yother JL, Briles DE;  
 XX WPI; 1999-105118/09.  
 XX N-PSDB; AAV84069.  
 XX DNA encoding truncated pneumococcal PspA protein - useful for producing  
 FT recombinant truncated protein.  
 XX  
 XX Disclosure; Fig 3A-D; 27pp; English.  
 XX  
 XX The present sequence represents a pneumococcal surface protein (PspA) of  
 CC Streptococcus pneumoniae Rxi. The specification describes truncated forms  
 CC of PspA which contain at least the N-terminal alpha-helical protective  
 CC epitope-containing region. The truncated PspA proteins can be used in  
 CC vaccines against pneumococcal infections  
 XX  
 SQ Sequence 648 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 648;  
 Best Local Similarity 98.0%; Pred. No. 2e-33;

Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 60  
 DB 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 281  
 QY 61 EENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEP 100  
 DB 282 EENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEP 321  
 RESULT 14  
 AAW92456  
 ID AAW92456 standard; protein; 653 AA.  
 XX  
 XX AAW92456;  
 XX  
 XX 21-APR-1999 (first entry)  
 XX  
 XX S. pneumoniae truncated pspA protein.  
 XX PspA; pneumococcal surface protein A; immunoprotection; detection; GMI;  
 KW solid phase immunosorbent assay; epitope; cell membrane anchor region;  
 KW cholera toxin B subunit; CTB; monosialoganglioside; fusion protein.  
 XX Streptococcus pneumoniae.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 620 /note= "in frame stop codon encoded by TAA"  
 FT Misc-difference 623 /note= "In frame stop codon encoded by TAA"  
 FT Misc-difference 628 /note= "In frame stop codon encoded by TAA"  
 FT Misc-difference 634 /note= "in frame stop codon encoded by TAA"  
 FT Misc-difference 636 /note= "In frame stop codon encoded by TGA"  
 XX  
 XX US5871943-A.  
 XX  
 XX 16-FEB-1999.  
 XX  
 XX 06-JUN-1995; 95US-00468718.  
 XX  
 XX 15-FEB-1991; 91US-00656773.  
 XX 12-FEB-1992; 92US-00835698.  
 XX 03-JUN-1993; 93US-00072068.  
 XX (UABR-) UAB RES FOUND.  
 XX  
 XX Briles DE, Yother JL;  
 XX WPI; 1999-166635/14.  
 XX N-PSDB; AAX02012.  
 XX Immunosorbent assay for pneumococcal surface protein A antigen or  
 PT antibody - for diagnosis of infection by Streptococcus pneumoniae.  
 XX  
 XX Claim 1; Fig 3A-C; 24pp; English.  
 XX  
 XX This sequence represents a truncated form of the Streptococcus pneumoniae  
 CC PspA protein which is used in a solid phase immunosorbent assay for  
 CC detecting a PspA (pneumococcal surface protein A) antibody and antigen.  
 CC This truncated protein contains the immunoprotective epitopes of the  
 CC complete protein (up to 90% of PspA but excludes the cell-membrane anchor  
 CC region) fused to the B subunit of cholera toxin (CTB) which is bound to  
 CC monosialoganglioside (GMI) coated on the substrate. The use of a fusion  
 CC between truncated PspA and cholera toxin B subunit (CTB) allows the  
 CC support to be coated without having to isolate PspA fragments since CTB  
 CC binds specifically to the GMI coating the solid support  
 XX  
 XX Sequence 653 AA;

Query Match 96.0%; Score 469.5; DB 2; Length 653;  
 Best Local Similarity 98.0%; Pred. No. 2e-33; Mismatches 1; Gaps 1;  
 Matches 98; Conservative 1; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLOSKLDKQAKLSKLELSKDIDELDAETAKLEDQLKAA 60  
 |||||:|||||  
 Db 223 LKEIDSESEDYAKGFRAPLOSKLDKQAKLSKLELSKDIDELDAETAKLEDQLKAA 281  
 |||||:|||||

QY 61 ENNNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100  
 |||||:|||||  
 Db 282 ENNNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321  
 |||||:|||||

## RESULT 15

AAR73912  
 ID AAR73912 standard; protein; 684 AA.

AC AAR73912;

DT 25-MAR-2003 (revised)

DT 05-DEC-1995 (first entry)

DE Streptococcus pneumoniae surface protein A.

XX Streptococcus pneumoniae; surface protein A; vaccine;

KW meningitis related homologous antigenic sequence; MRHAS; RV-1;

KW immunoassay; diagnosis; treatment; prophylactic; bacterial; viral.

XX Streptococcus pneumoniae.

PH Key Location/Qualifiers

FT Peptide 1..5

FT /label= sig\_peptide

PN WO9509232-A2.

XX 06-APR-1995.

XX 28-SEP-1994; 94WO-CA000516.

XX 28-SEP-1993; 93US-00127499.

XX (VALS/) VAN ALSTYNE D.

PA (SHAR/) SHARMA L R.

XX Van Alstyne D, Sharma LR;

PI WPI; 1995-147431/19.

XX New peptide(s) and corresp. antibodies for the treatment of meningitis -  
 PT the peptide(s) corresp. to homologous antigenic sites on bacterial and  
 PT viral agents and on chemokine(s), used for detecting and preventing  
 PT meningitis.

PS Claim 47; Fig 6/10; 98pp; English.

XX AAR73912 is the Streptococcus pneumoniae surface protein A. It contains  
 CC the meningitis related antigenic sequences (MRHAS) claimed in AAR73990  
 CC and AAR73902, which are recognised by a monoclonal antibody from the  
 CC hybridoma Rubella virus (RV)-1. The claimed MRHAS peptides may be used in  
 CC immunoassays to diagnose the presence of bacterial and/or viral  
 CC meningitis agents in a sample, or in prophylactic and therapeutic  
 CC meningitis treatments. The peptides may also be used as vaccines against  
 CC meningitis. NB: Identified by matching corresponding MRHAS peptides.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 684 AA;

Query Match

Best Local Similarity 96.0%; Score 469.5; DB 2; Length 684;

Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLOSKLDKQAKLSKLELSKDIDELDAETAKLEDQLKAA 60  
 |||||:|||||  
 Db 259 LKEIDSESEDYAKGFRAPLOSKLDKQAKLSKLELSKDIDELDAETAKLEDQLKAA 317  
 |||||:|||||

QY 61 ENNNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100  
 |||||:|||||

Db 318 ENNNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 357  
 |||||:|||||

Search completed: June 21, 2005, 10:10:13

Job time : 74.5918 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.6735 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-12  
Perfect score: 489  
Sequence: 1 LKEIDSESEDYAKGFRAP.....KKAELKTEADLKKAWEPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 489   | 100.0       | 100    | 4     | US-09-147-875A-12 |
| 2          | 469.5 | 96.0        | 99     | 2     | US-08-710-749-11  |
| 3          | 469.5 | 96.0        | 198    | 4     | US-08-529-055-61  |
| 4          | 469.5 | 96.0        | 619    | 1     | US-08-465-746-2   |
| 5          | 469.5 | 96.0        | 619    | 1     | US-08-214-164-2   |
| 6          | 469.5 | 96.0        | 619    | 2     | US-08-467-852A-3  |
| 7          | 469.5 | 96.0        | 619    | 2     | US-08-246-636-2   |
| 8          | 469.5 | 96.0        | 619    | 2     | US-08-247-491A-3  |
| 9          | 469.5 | 96.0        | 619    | 2     | US-08-319-795-2   |
| 10         | 469.5 | 96.0        | 619    | 2     | US-08-468-985-2   |
| 11         | 469.5 | 96.0        | 619    | 3     | US-08-312-949-2   |
| 12         | 469.5 | 96.0        | 648    | 1     | US-08-072-070-2   |
| 13         | 469.5 | 96.0        | 648    | 1     | US-08-469-434-2   |
| 14         | 469.5 | 96.0        | 648    | 1     | US-08-214-222-2   |
| 15         | 469.5 | 96.0        | 648    | 2     | US-08-467-852A-2  |
| 16         | 469.5 | 96.0        | 648    | 2     | US-08-468-718-2   |
| 17         | 469.5 | 96.0        | 648    | 2     | US-08-247-491A-2  |
| 18         | 469.5 | 96.0        | 648    | 3     | US-08-446-201-3   |
| 19         | 469.5 | 96.0        | 695    | 1     | US-08-127-499A-23 |
| 20         | 469.5 | 96.0        | 695    | 1     | US-08-482-847-23  |
| 21         | 469.5 | 96.0        | 8991   | 4     | US-08-714-741-32  |
| 22         | 462.5 | 94.6        | 99     | 2     | US-08-710-749-10  |
| 23         | 462.5 | 94.6        | 99     | 4     | US-09-147-875A-11 |
| 24         | 462.5 | 94.6        | 204    | 4     | US-08-529-055-51  |
| 25         | 457.5 | 93.6        | 288    | 3     | US-08-312-949-4   |
| 26         | 457.5 | 93.6        | 288    | 3     | US-08-446-201-4   |
| 27         | 447   | 91.4        | 289    | 1     | US-08-072-070-4   |

|    |       |      |      |   |                   |                    |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 447   | 91.4 | 289  | 1 | US-08-469-434-4   | Sequence 4, Appli  |
| 29 | 447   | 91.4 | 289  | 1 | US-08-214-222-4   | Sequence 4, Appli  |
| 30 | 447   | 91.4 | 289  | 1 | US-08-467-852A-5  | Sequence 5, Appli  |
| 31 | 447   | 91.4 | 289  | 2 | US-08-468-718-4   | Sequence 4, Appli  |
| 32 | 447   | 91.4 | 289  | 2 | US-08-247-491A-5  | Sequence 5, Appli  |
| 33 | 446.5 | 91.3 | 99   | 4 | US-09-147-875A-13 | Sequence 13, Appli |
| 34 | 442.5 | 90.5 | 99   | 2 | US-08-710-749-12  | Sequence 12, Appli |
| 35 | 442.5 | 90.5 | 195  | 4 | US-08-529-055-71  | Sequence 71, Appli |
| 36 | 426.5 | 87.2 | 1231 | 4 | US-08-714-741-47  | Sequence 41, Appli |
| 37 | 423.5 | 86.6 | 623  | 4 | US-08-714-741-41  | Sequence 47, Appli |
| 38 | 412   | 84.3 | 100  | 4 | US-09-147-875A-10 | Sequence 10, Appli |
| 39 | 407.5 | 83.3 | 99   | 2 | US-08-710-749-17  | Sequence 17, Appli |
| 40 | 402.5 | 82.3 | 170  | 4 | US-08-529-055-60  | Sequence 60, Appli |
| 41 | 402.5 | 82.3 | 181  | 4 | US-08-529-055-42  | Sequence 42, Appli |
| 42 | 402.5 | 82.3 | 864  | 4 | US-08-714-741-40  | Sequence 40, Appli |
| 43 | 399.5 | 81.7 | 99   | 4 | US-09-147-875A-16 | Sequence 16, Appli |
| 44 | 399.5 | 81.7 | 188  | 4 | US-08-529-055-59  | Sequence 59, Appli |
| 45 | 392.5 | 80.3 | 204  | 4 | US-08-529-055-58  | Sequence 58, Appli |

ALIGNMENTS

RESULT 1  
US-09-147-875A-12  
; Sequence 12, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-12

|                       |        |   |       |            |   |        |     |
|-----------------------|--------|---|-------|------------|---|--------|-----|
| Query Match           | 100.0% | Score   | 489   | DB         | 4 | Length | 100 |
| Best Local Similarity | 100.0% | Pred. No.   | 4e-37 |            |   |        |     |
| Matches               | 100    | Conservative  | 0     | Mismatches | 0 | Indels | 0   |
| Gaps                  | 0      |   |       |            |   |        |     |
| QY                    | 1      | LKEIDSESEDYAKGFRAPLQSKLDKAKLSKLELSKDIDELDAEIAKLEDQKAA | 60    |            |   |        |     |
| DB                    | 1      | LKEIDSESEDYAKGFRAPLQSKLDKAKLSKLELSKDIDELDAEIAKLEDQKAA | 60    |            |   |        |     |
| QY                    | 61     | EENNVEDYFKGLEKTIKAAKAELEKTEADLKKAWEPE                 | 100   |            |   |        |     |
| DB                    | 61     | EENNVEDYFKGLEKTIKAAKAELEKTEADLKKAWEPE                 | 100   |            |   |        |     |

RESULT 2  
US-08-710-749-11  
; Sequence 11, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: 530 Fifth Avenue  
; STREET: Curtis, Morris & Safford  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-11

Query Match 96.0%; Score 469.5; DB 2; Length 99;  
Best Local Similarity 98.0%; Pred. No. 2.2e-35;  
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKDIDELDAEIAKLQDLKAA 60  
Db 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKDIDELDAEIAKLQDLKAA 59  
Qy 61 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 100  
Db 60 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 99

## RESULT 3

US-08-529-055-61  
Sequence 61, Application US/08529055  
Patent No. 6592876  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: Pneumococcal Genes, Portions  
TITLE OF INVENTION: Thereof, Expression Products  
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
TITLE OF INVENTION: Portions and Products  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-61

Query Match 96.0%; Score 469.5; DB 4; Length 198;  
Best Local Similarity 98.0%; Pred. No. 5e-35;  
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKDIDELDAEIAKLQDLKAA 60  
Db 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKDIDELDAEIAKLQDLKAA 59  
Qy 61 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 100  
Db 60 EENNVDYFKGLEKTIAAKAELEKTEADLKKAVNEP 99

## RESULT 4

US-08-465-746-2  
Sequence 2, Application US/08465746  
Patent No. 5679768  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: Yother, Janet L.  
APPLICANT: McDaniel, Larry S.  
TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE  
TITLE OF INVENTION: PROTEIN A  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Shoemaker and Mattare, Ltd  
STREET: Suite 1203, 2001 Jefferson Davis Highway  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,746  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/048,896  
FILING DATE:  
APPLICATION NUMBER: US 07/656,773  
FILING DATE: 15-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/835,698  
FILING DATE: 12-FEB-1992  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 521-0378  
TELEX: LUKPAT WASHINGTON  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 619 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-465-746-2

Query Match 96.0%; Score 469.5; DB 1; Length 619;  
Best Local Similarity 98.0%; Pred. No. 1.9e-34;

```
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQL-KAA 281
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 5
US-08-214-164-2
; Sequence 2, Application US/08214164
; Patent No. 5728387
; GENERAL INFORMATION:
; APPLICANT: BRILES, DAVID E.
; APPLICANT: YOTHEER, JANET L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-MAR-1994
; APPLICATION NUMBER: US/08/214,164
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W.
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 6102-137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0813
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-214-164-2

Query Match 96.0%; Score 469.5; DB 1; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQL-KAA 281
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 6
US-08-467-852A-3
; Sequence 3, Application US/08467852A
; Patent No. 5856170
```

```
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTHEER, Janet L.
; APPLICANT: MCDANIEL, Larry S.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,852A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-467-852A-3

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQLKAA 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLSKLEELSDKIDELDAETAKLEDQL-KAA 281
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
Db 282 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 7
US-08-246-636-2
; Sequence 2, Application US/08246636
; Patent No. 5985141
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yotther, Janet L
; APPLICANT: McDaniel, Larry S
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,636  
;; FILING DATE: 20-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/656,773  
;; FILING DATE: 15-FEB-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/835,698  
;; FILING DATE: 12-FEB-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/048,896  
;; FILING DATE: 20-APR-1993  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 415-0810  
;; TELEX: LUKPAT WASHINGTON  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 619 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-246-636-2

Query Match 96.0%; Score 469.5; DB 2; Length 619;  
Best Local Similarity 98.0%; Pred. No. 1.9e-34;  
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
  
Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLKAA 60  
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSKIDELDAEIAKLEDLKAA 281  
  
Qy 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 100  
Db 282 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 8  
US-08-247-491A-3  
; Sequence 3, Application US/08247491A  
; Patent No. 5965400  
; GENERAL INFORMATION:  
; APPLICANT: BRILES, David E.  
; APPLICANT: YOTHER, Janet L.  
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10151  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/247,491A  
; FILING DATE: 23-JUN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOWALSKI, Thomas J.  
; REGISTRATION NUMBER: 32,147  
; REFERENCE/DOCKET NUMBER: 454312-2041  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-588-0800  
; TELEFAX: 212-588-0500  
; INFORMATION FOR SEQ ID NO: 3:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 619 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: n/a  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: amino acid  
US-08-247-491A-3  
  
Query Match 96.0%; Score 469.5; DB 2; Length 619;  
Best Local Similarity 98.0%; Pred. No. 1.9e-34;  
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
  
Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEDLKAA 60  
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSKIDELDAEIAKLEDLKAA 281  
  
Qy 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 100  
Db 282 EENNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 9  
US-08-319-795-2  
; Sequence 2, Application US/08319795  
; Patent No. 5980909  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Yother, Janet L.  
; APPLICANT: McDaniel, Larry S  
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface  
; TITLE OF INVENTION: Protein A  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheomaker and Mattare, Ltd.  
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/319,795  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/246,636  
; FILING DATE: 20-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/048,896  
; FILING DATE: 20-APR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/835,698  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/656,773  
; FILING DATE: 15-FEB-1991  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 415-0810  
; TELEFAX: (703) 415-0813  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 619 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-319-795-2

Query Match 96.0%; Score 469.5; DB 2; Length 619;



```
Best Local Similarity 98.0%; Pred. No. 1.9e-34; Mismatches 1; Indels 0; Gaps 1;
Matches 98; Conservative 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQ-KAA 281

QY 61 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 100
Db 282 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 321

RESULT 10
US-08-468-985-2
; Sequence 2, Application US/08468985
; Patent No. 5997882
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
; TITLE OF INVENTION: Protein A
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,985
; FILING DATE:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/319,795
; FILING DATE:
; APPLICATION NUMBER: US 08/246,636
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,896
; FILING DATE: 20-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-985-2

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 99; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQ-KAA 281

QY 61 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 100
Db 282 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 321

RESULT 11
US-08-312-949-2
; Sequence 2, Application US/08312949
; Patent No. 6027734
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
; TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,949
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-312-949-2

Query Match 96.0%; Score 469.5; DB 3; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQ-KAA 281

QY 61 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 100
Db 282 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 321

RESULT 12
US-08-072-070-2
; Sequence 2, Application US/08072070
; Patent No. 5476929
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; 
```

```
QY 61 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 100
Db 282 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 321

RESULT 11
US-08-312-949-2
; Sequence 2, Application US/08312949
; Patent No. 6027734
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
; TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,949
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-312-949-2

Query Match 96.0%; Score 469.5; DB 3; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.9e-34;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQKAA 60
Db 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQ-KAA 281

QY 61 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 100
Db 282 EENNVDYFKGGLKTIKAAKAELEKTEADLKKAVNEPE 321

RESULT 12
US-08-072-070-2
; Sequence 2, Application US/08072070
; Patent No. 5476929
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; 
```

STREET: Suite 1203, 2001 Jefferson Davis Highway  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/072,070  
FILING DATE: 19930603  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/835,698  
FILING DATE: 12-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/656,773  
FILING DATE: 15-FEB-1991  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 521-0378  
TELEX: LUKPAT WASHINGTON  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 648 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-072-070-2

Query Match 96.0%; Score 469.5; DB 1; Length 648;  
Best Local Similarity 98.0%; Pred. No. 2e-34; Mismatches 0; Indels 1; Gaps 1;

Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 60  
|||||  
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 281  
|||||

Qy 61 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 100  
|||||  
Db 282 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321  
|||||

RESULT 13  
US-08-469-434-2  
Sequence 2, Application US/08469434  
Patent No. 5753463  
GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: Yother, Janet L  
APPLICANT: McDaniel, Larry S  
TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Shoemaker and Mattare, Ltd  
STREET: Suite 1203, 2001 Jefferson Davis Highway  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,434  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/072,065  
FILING DATE: 03 JUNE 1993  
APPLICATION NUMBER: US/07/835,698  
FILING DATE: 12-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/656,773  
FILING DATE: 15-FEB-1991  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 521-0378  
TELEX: LUKPAT WASHINGTON  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 648 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-434-2

Query Match 96.0%; Score 469.5; DB 1; Length 648;  
Best Local Similarity 98.0%; Pred. No. 2e-34; Mismatches 1; Gaps 1;  
Matches 98; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 60  
|||||  
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAA 281  
|||||

Qy 61 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 100  
|||||  
Db 282 EENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321  
|||||

RESULT 14  
US-08-214-222-2  
Sequence 2, Application US/08214222  
Patent No. 5804193  
GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: Yother, Janet L  
TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Shoemaker and Mattare, Ltd  
STREET: Suite 1203, 2001 Jefferson Davis Highway  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/214,222  
FILING DATE: 17-MAR-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/835,698  
FILING DATE: 12-FEB-1992  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 521-0378  
TELEX: LUKPAT WASHINGTON  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 648 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-214-222-2

Query Match 96.0%; Score 469.5; DB 1; Length 648;  
Best Local Similarity 98.0%; Pred. No. 2e-34;  
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLELDQKKAA 60  
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLELDQKKAA 281  
QY 61 EENNVEDYFKEGLEKTIAAKKALEKTEADLKKAVNEPE 100  
DB 282 EENNVEDYFKEGLEKTIAAKKALEKTEADLKKAVNEPE 321

## RESULT 15

US-08-467-852A-2  
; Sequence 2, Application US/08467852A  
; Patent No. 5856170  
; GENERAL INFORMATION:  
; APPLICANT: BRILES, David E.  
; APPLICANT: YOTHER, Janet L.  
; APPLICANT: MCDANIEL, Larry S.  
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10151

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,852A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: KOWALSKI, Thomas J.  
; REGISTRATION NUMBER: 32,147  
; REFERENCE/DOCKET NUMBER: 454312-2064  
; TELEPHONE: 212-588-0800  
; TELEFAX: 212-588-0500  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 648 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

## US-08-467-852A-2

Query Match 96.0%; Score 469.5; DB 2; Length 648;  
Best Local Similarity 98.0%; Pred. No. 2e-34;  
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLELDQKKAA 60  
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLELDQKKAA 281  
QY 61 EENNVEDYFKEGLEKTIAAKKALEKTEADLKKAVNEPE 100  
DB 282 EENNVEDYFKEGLEKTIAAKKALEKTEADLKKAVNEPE 321

Search completed: June 21, 2005, 10:25:19  
Job time : 18.6735 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.8776 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-12

Perfect score: 489

Sequence: 1 LKEIDSESEDYAKEGFRAP.....KKAELKTEADLKAVNEPE 100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 489   | 100.0       | 100    | 15    | US-10-674-755-12     |
| 2          | 469.5 | 96.0        | 198    | 15    | US-10-299-636-76     |
| 3          | 469.5 | 96.0        | 354    | 15    | US-10-299-636-105    |
| 4          | 469.5 | 96.0        | 588    | 15    | US-10-299-636-96     |
| 5          | 469.5 | 96.0        | 619    | 10    | US-09-882-774-1      |
| 6          | 469.5 | 96.0        | 619    | 15    | US-10-282-122A-73702 |
| 7          | 469.5 | 96.0        | 619    | 16    | US-10-414-532-72     |
| 8          | 462.5 | 94.6        | 99     | 15    | US-10-674-755-11     |
| 9          | 462.5 | 94.5        | 204    | 15    | US-10-299-636-66     |
| 10         | 446.5 | 91.3        | 99     | 15    | US-10-674-755-13     |
| 11         | 442.5 | 90.5        | 195    | 15    | US-10-299-636-86     |

|    |       |      |     |    |                    |                   |
|----|-------|------|-----|----|--------------------|-------------------|
| 12 | 412   | 84.3 | 100 | 15 | US-10-674-755-10   | Sequence 10, Appl |
| 13 | 402.5 | 82.3 | 170 | 15 | US-10-299-636-75   | Sequence 75, Appl |
| 14 | 402.5 | 82.3 | 181 | 15 | US-10-299-636-57   | Sequence 57, Appl |
| 15 | 402.5 | 82.3 | 643 | 15 | US-10-299-636-95   | Sequence 95, Appl |
| 16 | 402.5 | 82.3 | 670 | 9  | US-09-748-875-63   | Sequence 63, Appl |
| 17 | 402.5 | 82.3 | 670 | 10 | US-09-298-523B-63  | Sequence 63, Appl |
| 18 | 402.5 | 82.3 | 690 | 9  | US-09-748-875-61   | Sequence 61, Appl |
| 19 | 402.5 | 82.3 | 690 | 10 | US-09-298-523B-61  | Sequence 61, Appl |
| 20 | 402.5 | 82.3 | 691 | 9  | US-09-748-875-1    | Sequence 1, Appl  |
| 21 | 402.5 | 82.3 | 691 | 10 | US-09-298-523B-1   | Sequence 1, Appl  |
| 22 | 402.5 | 82.3 | 701 | 9  | US-09-748-875-62   | Sequence 62, Appl |
| 23 | 402.5 | 82.3 | 701 | 10 | US-09-298-523B-62  | Sequence 62, Appl |
| 24 | 402.5 | 82.3 | 707 | 9  | US-09-748-875-2    | Sequence 2, Appl  |
| 25 | 402.5 | 82.3 | 707 | 10 | US-09-298-523B-2   | Sequence 2, Appl  |
| 26 | 402.5 | 82.3 | 711 | 9  | US-09-748-875-3    | Sequence 3, Appl  |
| 27 | 402.5 | 82.3 | 711 | 10 | US-09-298-523B-3   | Sequence 3, Appl  |
| 28 | 402.5 | 82.3 | 739 | 17 | US-10-732-923-3294 | Sequence 3294, Ap |
| 29 | 402.5 | 82.3 | 929 | 9  | US-09-748-875-60   | Sequence 60, Appl |
| 30 | 402.5 | 82.3 | 929 | 10 | US-09-298-523B-60  | Sequence 60, Appl |
| 31 | 402.5 | 82.3 | 929 | 15 | US-10-299-636-94   | Sequence 94, Appl |
| 32 | 399.5 | 81.7 | 99  | 15 | US-10-674-755-16   | Sequence 16, Appl |
| 33 | 399.5 | 81.7 | 188 | 15 | US-10-299-636-74   | Sequence 74, Appl |
| 34 | 392.5 | 80.3 | 204 | 15 | US-10-299-636-73   | Sequence 73, Appl |
| 35 | 390.5 | 79.9 | 141 | 14 | US-10-254-995-2    | Sequence 2, Appl  |
| 36 | 390.5 | 79.9 | 589 | 9  | US-09-748-875-14   | Sequence 14, Appl |
| 37 | 390.5 | 79.9 | 589 | 10 | US-09-298-523B-14  | Sequence 14, Appl |
| 38 | 390.5 | 79.9 | 589 | 15 | US-10-299-636-97   | Sequence 97, Appl |
| 39 | 387.5 | 79.2 | 99  | 15 | US-10-674-755-15   | Sequence 15, Appl |
| 40 | 375.5 | 76.8 | 206 | 15 | US-10-299-636-69   | Sequence 69, Appl |
| 41 | 372.5 | 76.2 | 99  | 15 | US-10-674-755-14   | Sequence 14, Appl |
| 42 | 346   | 70.8 | 100 | 15 | US-10-674-755-2    | Sequence 2, Appl  |
| 43 | 339   | 69.3 | 100 | 15 | US-10-674-755-3    | Sequence 3, Appl  |
| 44 | 334.5 | 68.4 | 73  | 9  | US-09-027-956-8    | Sequence 8, Appl  |
| 45 | 334   | 68.3 | 194 | 15 | US-10-299-636-79   | Sequence 79, Appl |

## ALIGNMENTS

## RESULT 1

```
US-10-674-755-12
; Sequence 12, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-12
```

Query Match 100.0%; Score 489; DB 15; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.6e-31;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESEDYAKEGFRAPLQSKLDAKQAKLSKLELSKDIDELDAETAKLEDQLKAA 60

Db 1 LKEIDSESEDYAKEGFRAPLQSKLDAKQAKLSKLELSKDIDELDAETAKLEDQLKAA 60

QY 61 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 100

Db 61 EENNVEDYFKEGLEKTIKAAKAELEKTEADLKAVNEPE 100

## RESULT 2

```
US-10-299-636-76
; Sequence 76, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-76

Query Match          96.0%; Score 469.5; DB 15; Length 198;
Best Local Similarity 98.0%; Pred. No. 1.9e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60

Qy 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
    |||||
Db 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100

Qy 60 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 99
    |||||

RESULT 3
US-10-299-636-105
; Sequence 105, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-105

Query Match          96.0%; Score 469.5; DB 15; Length 354;
Best Local Similarity 98.0%; Pred. No. 3.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60

Qy 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
    |||||
Db 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100

Qy 60 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 99
    |||||

RESULT 5
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
US-09-882-774-1

Query Match          96.0%; Score 469.5; DB 15; Length 588;
Best Local Similarity 98.0%; Pred. No. 6.3e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 60
    |||||
Db 192 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLELSDKIDELDAEIAKLEDLQAKAA 250

Qy 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
    |||||
Db 251 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 290

RESULT 5
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
US-09-882-774-1
```

```
; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

Query Match          96.0%; Score 469.5; DB 10; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQ-L-KAA 281

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 100
DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 6
US-10-282-122A-73702
; Sequence 73702, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73702
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73702

Query Match          96.0%; Score 469.5; DB 15; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQ-L-KAA 281

; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

Query Match          96.0%; Score 469.5; DB 10; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 100
DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 7
US-10-414-532-72
; Sequence 72, Application US/10414532
; Publication No. US20040101531A1
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 72
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-414-532-72

Query Match          96.0%; Score 469.5; DB 16; Length 619;
Best Local Similarity 98.0%; Pred. No. 6.6e-29;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQ-L-KAA 281

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 100
DB 282 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 321

RESULT 8
US-10-674-755-11
; Sequence 11, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-11

Query Match          94.6%; Score 462.5; DB 15; Length 99;
Best Local Similarity 96.0%; Pred. No. 3.2e-29;
Matches 96; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 1 LKEIDSESEDYAKGFRAPLQSKLDAAKQAKLSKLEELSDKIDELDAETAKLEDQ-L-KAA 59

QY 61 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 100
DB 60 EENNVEDYFKGLEKTIAAKKAELKTEADLKAVNEPE 99
```

```

Db      60 EENNVEDYSTGLEKTIAAKTELEKTEADLKAVNEPE 99

RESULT 9
US-10-299-636-66
; Sequence 66, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-66

Query Match      94.6%; Score 462.5; DB 15; Length 204;
Best Local Similarity 96.0%; Pred. No. 7e-29;
Matches 96; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 59

Qy 61 EENNVDYFKGELKTIAAKAELEKTEADLKAVNEPE 100
Db 60 EENNVDYFKGELKTIAAKAELEKTEADLKAVNEPE 99

RESULT 10
US-10-674-755-13
; Sequence 13, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-13

Query Match      91.3%; Score 446.5; DB 15; Length 99;
Best Local Similarity 94.0%; Pred. No. 5.8e-28;
Matches 94; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 59

Qy 61 EENNVDYFKGELKTIAAKAELEKTEADLKAVNEPE 100
Db 60 EENNVDYFKGELKTIAAKAELEKTEADLKAVNEPE 99

RESULT 11
US-10-299-636-86
; Sequence 86, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-86

Query Match      90.5%; Score 442.5; DB 15; Length 195;
Best Local Similarity 93.0%; Pred. No. 2.5e-27;
Matches 93; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 59

Qy 61 EENNVDYFKGELKTIAAKAELEKTEADLKAVNEPE 100
Db 60 EENNVEDYSTGLEKTIAAKTELEKTEADLKAVNEPE 99

RESULT 12
US-10-674-755-10
; Sequence 10, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-10

Query Match      84.3%; Score 412; DB 15; Length 100;
Best Local Similarity 87.0%; Pred. No. 3e-25;
Matches 87; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
Db 1 LKEIDSESDYAKGFRAPLQSKLDKQAKLSKLELSKIDELDAEIAKLEDLQKAA 60
```



QY 61 EENNVEDYFKEGLEKTIAAKKALEKTEADLKAVNEPE 100  
DB 61 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 100

## RESULT 13

US-10-299-636-75  
; Sequence 75, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 75  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-75

Query Match 82.3%; Score 402.5; DB 15; Length 170;  
Best Local Similarity 86.0%; Pred. No. 3e-24;  
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLOSKLDKAKQAKLSKLEELSDKIDELDAETAKLEDQKKAA 60  
DB 1 LKEIDSESDYKLEGLRAPLOSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKD-A 59

QY 61 EENNVEDYFKEGLEKTIAAKKALEKTEADLKAVNEPE 100  
DB 60 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 99

## RESULT 14

US-10-299-636-57  
; Sequence 57, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 57  
; LENGTH: 161  
; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-57

Query Match 82.3%; Score 402.5; DB 15; Length 181;  
Best Local Similarity 86.0%; Pred. No. 3.2e-24;  
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLOSKLDKAKQAKLSKLEELSDKIDELDAETAKLEDQKKAA 60  
DB 1 LKEIDSESDYKLEGLRAPLOSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKD-A 59  
QY 61 EENNVEDYFKEGLEKTIAAKKALEKTEADLKAVNEPE 100  
DB 60 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 99

## RESULT 15

US-10-299-636-95  
; Sequence 95, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 95  
; LENGTH: 643  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-95

Query Match 82.3%; Score 402.5; DB 15; Length 643;  
Best Local Similarity 86.0%; Pred. No. 1.3e-23;  
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLOSKLDKAKQAKLSKLEELSDKIDELDAETAKLEDQKKAA 60  
DB 245 LKEIDSESDYKLEGLRAPLOSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKD-A 303

QY 61 EENNVEDYFKEGLEKTIAAKKALEKTEADLKAVNEPE 100  
DB 304 EGNNVVEAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 343

Search completed: June 21, 2005, 11:18:34  
Job time : 64.8775 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 10 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-12  
Perfect score: 489  
Sequence: 1 LKEIDSESEDYAKGFRAP.....KKAELKTEADLKAVNEPE 100

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 469.5 | 96.0        | 619    | 2 A97887 | surface protein ps |
| 2          | 469.5 | 96.0        | 619    | 2 A41971 | surface protein ps |
| 3          | 144   | 29.4        | 744    | 2 P95013 | pneumococcal surfa |
| 4          | 120   | 24.5        | 161    | 2 S48396 | tropomyosin TPM2 - |
| 5          | 118   | 24.1        | 501    | 2 A44643 | M protein precurs  |
| 6          | 116.5 | 23.8        | 886    | 2 H69378 | conserved hypotet  |
| 7          | 116   | 23.7        | 3488   | 2 T34418 | hypothetical prote |
| 8          | 110   | 22.5        | 233    | 2 S70531 | bbk2.11 protein pr |
| 9          | 110   | 22.5        | 852    | 2 D72230 | conserved hypotet  |
| 10         | 109   | 22.3        | 230    | 2 I40287 | outer surface prot |
| 11         | 109   | 22.3        | 880    | 2 F75103 | conserved hypotet  |
| 12         | 108.5 | 22.2        | 281    | 2 F75216 | hypothetical prote |
| 13         | 107.5 | 22.0        | 229    | 2 S70532 | outer surface prot |
| 14         | 107.5 | 22.0        | 764    | 2 T05409 | hypothetical prote |
| 15         | 105   | 21.5        | 1312   | 2 T30845 | probable DNA repai |
| 16         | 104.5 | 21.4        | 1837   | 2 T41023 | probable nuclear p |
| 17         | 104   | 21.3        | 1006   | 2 C70445 | Arpase subunit of  |
| 18         | 103   | 21.1        | 2401   | 2 T28676 | rhostry protein -  |
| 19         | 102   | 20.9        | 388    | 2 A46173 | Mrp4 protein - Str |
| 20         | 102   | 20.9        | 388    | 2 S52536 | fcrA 15 protein -  |
| 21         | 102   | 20.9        | 405    | 2 A33939 | Pc gamma (19G) rec |
| 22         | 102   | 20.9        | 1177   | 2 B75150 | chromosome segrega |
| 23         | 101   | 20.7        | 372    | 2 S23326 | gene ML2.2 protein |
| 24         | 100   | 20.4        | 415    | 2 S35760 | fcrA protein precu |
| 25         | 100   | 20.4        | 1091   | 2 T34107 | hypothetical prote |
| 26         | 100   | 20.4        | 1319   | 2 A28313 | glued protein - fr |
| 27         | 100   | 20.4        | 3187   | 2 JCS937 | 364K Golgi complex |
| 28         | 99.5  | 20.3        | 166    | 2 S73342 | hypothetical prote |
| 29         | 99.5  | 20.3        | 473    | 2 F70031 | cell wall-binding  |

```
fcra protein precu
hypothetical prote
P115 homolog - Met
hypothetical prote
IGA receptor - Str
M2 protein precurs
hypothetical prote
epidermal growth f
NF-180 - sea lamp
myosin heavy chain
myosin heavy chain
hypothetical prote
hypothetical prote
fKBP-type peptidyl
radixin - human
radixin - pig

30 99 20.2 387 2 S57834
31 99 20.2 1053 2 T51375
32 99 20.2 1169 2 A64505
33 98 20.0 399 2 E71169
34 98 20.0 402 2 S37046
35 98 20.0 407 2 S23325
36 98 20.0 670 2 P84899
37 97.5 19.9 896 2 S43074
38 97.5 19.9 1110 2 I51116
39 97.5 19.9 1938 1 A40997
40 97.5 19.9 2139 2 T18296
41 97.5 19.9 3450 2 T26963
42 97.5 19.9 3461 2 T26964
43 97 19.8 431 2 A97225
44 97 19.8 583 1 A46127
45 97 19.8 583 1 S39805

ALIGNMENTS

RESULT 1
A97887
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A97887
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; Mc
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97887
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <XUR>
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:AE007317; PIDN:AAK98925.1; PID:gt1
C:Genetics:
A:Gene: pspA

Query Match 96.0%; Score 469.5; DB 2; Length 619;
Best Local Similarity 98.0%; Pred. No. 1.2e-23;
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 LKEIDSESEDYAKGFRAPLOSKLDKAKLSKLELSKDIDELDAETAKLEDQLKAA 60
|||||
Db 223 LKEIDSESEDYAKGFRAPLOSKLDKAKLSKLELSKDIDELDAETAKLEDQ-L-KAA 281
|||||

QY 61 EENNNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 100
|||||
Db 282 EENNNVEDYFKEGLEKTIAAKKAELKTEADLKAVNEPE 321
|||||

RESULT 2
A41971
surface protein pspA precursor - Streptococcus pneumoniae
N:Alternate names: pneumococcal surface protein A
C:Species: Streptococcus pneumoniae
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41971; A60282; A33134
R:Yother, J.; Briles, D.E.
J. Bacteriol. 174, 601-609, 1992
A>Title: Structural properties and evolutionary relationships of PspA, a surface protein
A:Reference number: A41971; MUID:92105030; PMID:1729249
A:Accession: A41971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <YOT>
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:M74122; NID:q153840; PIDN:AAA2701f
A>Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIIP:75636)
R:Talkington, D.F.; Crammins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.
```



Query Match 24.1%; Score 118; DB 2; Length 501;  
 Best Local Similarity 37.0%; Pred. No. 0.75;  
 Matches 37; Conservative 14; Mismatches 35; Indels 14; Gaps 4;

QY 3 EIDSESESDYAK-EGFRAPLQSKLDAKQAKSLKLELSKDIDELDAEIAKLELDQKAAE 61  
 DB 125 ETEKEVEDYKLVDEKASLDKIESAN---SOLEFNSQISLVQAANLNEQIEKLS 181

QY 62 ENNVEDYFKGLE-----KTTAAKAAELE-----KTEAD 91  
 DB 182 EKNKAEQSQNEQLEFKNQNIADLIGKAAELMKLAKAEED 221

RESULT 6  
 H69378  
 conserved hypothetical protein AF1032 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Accession: H69378  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: H69378  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-886 <KLE>  
 A:Cross-references: UNIPROT:O92230; GB:AE001032; GB:AE000782; NID:g26899355; PIDN:AA89021  
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 23.8%; Score 116.5; DB 2; Length 886;  
 Best Local Similarity 28.1%; Pred. No. 1.6;  
 Matches 38; Conservative 22; Mismatches 30; Indels 45; Gaps 5;

QY 1 LKEIDSESESDYAK-EGFRAPLQSKLDAKQAKSLKLELSKDIDELDAEI----- 49  
 DB 303 LRDEKREG-DLTREA--AGIQAKKAEEDNSKLEIEITKRIEELERELERPEKSHRLLE 359

QY 50 -----AKLE-----DQKAAEENNVEDYFKGLEKTIKAAKKA 83  
 DB 360 TLKPKMDRMQGIKAKLEENLTPDKVEMVLLSKAKEBEKEITKLLK-----LIAKKS 414

QY 84 ELEKTEADLKAAVNE 98  
 DB 415 SLKTRGAQLKKAVEE 429

RESULT 7  
 T34418  
 hypothetical protein F12F3.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34418  
 R:Fulton, B.; Wohlmann, P.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: The sequence of C. elegans cosmid F12F3.  
 A:Reference number: Z21521  
 A:Accession: T34418  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-3488 <FUL>  
 A:Cross-references: EMBL:U80022; PIDN:AA25885.1; GSPDB:GN00023; CESP:F12F3.3  
 A:Experimental source: strain Bristol N2; clone F12F3  
 C:Genetics:  
 A:Gene: CESP:F12F3.3  
 A:Map position: 5  
 A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 23.7%; Score 116; DB 2; Length 3488;

Best Local Similarity 37.6%; Pred. No. 6.5;  
 Matches 47; Conservative 18; Mismatches 26; Indels 34; Gaps 9;

QY 2 KEIDSESESDYAK-EGFRAPLQSKLDAKQAKSLKLELSKDIDELDAEI 49  
 DB 1009 KETDEKLDAIAAKTQEADEKSKLDA-QEKIKKVSDDAARKEKELNDKL-KLESEI 1066

QY 50 A-----KLEED--OLKAAE-----ENNVEDYFK-----EGLEKTIKAAKAELEKTEA 90  
 DB 1067 ATKQASADKLLEEQAKKAAEVEAAKQKQEKQKLDTEAASKAAAEKLELEK-QA 1125

QY 91 DLKKA 95  
 DB 1126 QIKKA 1130

RESULT 8  
 S70531  
 bbk2.11 protein precursor - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S70531  
 R:Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgard  
 Mol. Microbiol. 18, 507-520, 1995  
 A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer  
 A:Reference number: S70531; MUID:96342380; PMID:8748034  
 A:Accession: S70531  
 A>Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-233 <AKI>  
 A:Cross-references: UNIPROT:Q44739; EMBL:U30617; NID:g3309515; PIDN:AA46421.1; PID:g119  
 C:Superfamily: outer surface protein F ospF  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-233/Product: bbk2.11 protein #status predicted <MAT>

Query Match 22.5%; Score 110; DB 2; Length 233;  
 Best Local Similarity 31.0%; Pred. No. 1.2;  
 Matches 36; Conservative 24; Mismatches 32; Indels 24; Gaps 6;

QY 6 ESESDYAK-----EGFRAPLQ-----SKLDAKQAK--LSKLEELSDKIDELDAE-- 48  
 DB 37 ESEBQNVKTEQIKKQVEGFLEILETKDLSKLEDEKTEIEKQIQELKNKIEKLSKKT 96

QY 49 -----TAKLEDQKAAE--ENNVEDYFKGLEKTIKAAKAELEKTEADLKAAVNE 98  
 DB 97 SIETSYSEBEKINKIEKLKGGLEDKPKYE--LEESLAKKKGKRRKALQEAQKQFEE 151

RESULT 9  
 D72230  
 conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: D72230  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
 C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
 A:Reference number: A72200; MUID:99287316; PMID:10360571  
 A:Accession: D72230  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-852 <ARN>  
 A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AA03670  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM1636  
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.5%; Score 110; DB 2; Length 852;  
 Best Local Similarity 30.0%; Pred. No. 4.1;  
 Matches 27; Conservative 22; Mismatches 29; Indels 12; Gaps 2;



C:Accession: T05409  
 R:Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; N  
 submitted to the Protein Sequence Database, February 1998  
 A:Reference number: Z15414  
 A:Accession: T05409  
 A:Molecule type: DNA  
 A:Residues: 1-764 <BEV>  
 A:Cross-references: UNIPROT:O49371; EMBL:AL021811  
 A:Experimental source: cultivar Columbia; BAC clone F10M6  
 C:Genetics:  
 A:Map position: 4  
 A:Note: F10M6.170

Query Match 22.0%; Score 107.5; DB 2; Length 764;  
 Best Local Similarity 29.8%; Pred. No. 5.4;  
 Matches 36; Conservative 23; Mismatches 31; Indels 31; Gaps 4;  
 QY 2 KEIDSESDYAKGFRAPLQSKDAQAKSL-EELSDKIDELDAEIAKLEDO----- 55  
 DB 163 REIEELKHKLRDEREERAAQSSLTUKEEELKQOEIANRKEVSMALSEFESKQLLS 222  
 QY 56 -----LKKAAEENNVEDYFKGLEKTIAAKKALEK---TEADLKKAVN 97  
 DB 223 KANEVVRQEGEYIALQRALEK-----EELEISKATKKLEQEKLETEANLKKQTE 275  
 QY 98 E 98  
 DB 276 E 276

RESULT 15  
 T30845  
 probable DNA repair protein RAD50 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T30845  
 R:Kim, K.K.; Daud, A.I.; Wong, S.C.; Pajak, L.; Tsai, S.C.; Wang, H.; Henzel, W.J.; Fiel  
 J. Biol. Chem. 271, 29255-29284, 1996  
 A:Title: Mouse RAD50 has limited epitopic homology to p53 and is expressed in the adult  
 A:Reference number: Z20899; MUID:97067183; PMID:8910585  
 A:Accession: T30845  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1312 <KIM>  
 A:Cross-references: UNIPROT:P70388; EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AAC528  
 C:Genetics:  
 A:Gene: RAD50  
 A:Map position: 11  
 C:Superfamily: RAD50 protein  
 C:Keywords: DNA repair

Query Match 21.5%; Score 105; DB 2; Length 1312;  
 Best Local Similarity 35.1%; Pred. No. 13;  
 Matches 27; Conservative 20; Mismatches 20; Indels 10; Gaps 2;  
 QY 22 QSKLDKAKLSKLELSKDIDELDAEIAKLEDLQKAAEENNVEDYFKGLEKTIAAK 81  
 DB 454 QSELHVRVRSGLQEGSDRILELDQELTKAERLSK-AEKNSSIE-----TLKAE 503  
 QY 82 KAELEKTEADLKKAUNE 98  
 DB 504 VMSLQNEKADLDRSLRK 520

Search completed: June 21, 2005, 10:11:58  
 Job time : 11 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.9388 Seconds  
(without alignments)

826.751 Million cell updates/sec

Title: US-10-674-755-12

Perfect score: 489

Sequence: 1 LKEIDSESEDYAKGFRAP.....KKAELEKTEADLKKAYPE 100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | ID     | Description |
|------------|-------|---------------|--------|--------|-------------|
| 1          | 469.5 | 96.0          | 619    | Q54972 | streptococc |
| 2          | 469.5 | 96.0          | 619    | Q8DR10 | streptococc |
| 3          | 462.5 | 94.6          | 417    | Q9LAY3 | streptococc |
| 4          | 447.5 | 91.5          | 415    | Q9LAY1 | streptococc |
| 5          | 402.5 | 82.3          | 739    | Q9ROT4 | streptococc |
| 6          | 402.5 | 82.3          | 820    | Q9ROT1 | streptococc |
| 7          | 402.5 | 82.3          | 929    | Q9KK19 | streptococc |
| 8          | 402.5 | 82.3          | 929    | Q9ZAY5 | streptococc |
| 9          | 389.5 | 79.7          | 437    | Q9LAY4 | streptococc |
| 10         | 387.5 | 79.2          | 395    | Q9LAY2 | streptococc |
| 11         | 387.5 | 79.2          | 408    | Q9LAY0 | streptococc |
| 12         | 383.5 | 78.4          | 99     | Q8KQK4 | streptococc |
| 13         | 383.5 | 78.4          | 249    | Q9LS75 | streptococc |
| 14         | 376.5 | 77.0          | 224    | Q8GNS8 | streptococc |
| 15         | 370.5 | 75.8          | 426    | Q9LAY5 | streptococc |
| 16         | 368   | 75.3          | 869    | Q9KK27 | streptococc |
| 17         | 357   | 73.0          | 225    | Q9LS91 | streptococc |
| 18         | 351   | 71.8          | 222    | Q9LS77 | streptococc |
| 19         | 351   | 71.8          | 262    | Q9LS76 | streptococc |
| 20         | 351   | 71.8          | 415    | Q9LAY7 | streptococc |
| 21         | 347   | 71.0          | 246    | Q9LS78 | streptococc |
| 22         | 344   | 70.3          | 416    | Q9LAY8 | streptococc |
| 23         | 342   | 69.9          | 255    | Q9LS81 | streptococc |
| 24         | 342   | 69.9          | 255    | Q9LS86 | streptococc |
| 25         | 340   | 69.5          | 406    | Q9LAZ0 | streptococc |
| 26         | 335   | 68.5          | 393    | Q9LAZ3 | streptococc |
| 27         | 334   | 68.3          | 394    | Q9LAY6 | streptococc |
| 28         | 334   | 68.3          | 395    | Q9LAZ1 | streptococc |
| 29         | 329   | 67.3          | 340    | Q8KQK5 | streptococc |
| 30         | 328   | 67.1          | 194    | Q9LSB5 | streptococc |
| 31         | 328   | 67.1          | 218    | Q6UEB2 | streptococc |

32 328 67.1 233 2 Q9LS68 streptococc  
33 328 67.1 236 2 Q9LS69 streptococc  
34 328 67.1 243 2 Q9LS64 streptococc  
35 328 67.1 243 2 Q9LS67 streptococc  
36 328 67.1 244 2 Q9LS65 streptococc  
37 328 67.1 247 2 Q9LS66 streptococc  
38 328 67.1 249 2 Q9LS70 streptococc  
39 328 67.1 254 2 Q9LS63 streptococc  
40 328 67.1 401 2 Q9LAZ2 streptococc  
41 326 66.7 207 2 Q8GNS9 streptococc  
42 320 65.4 237 2 Q9LS92 streptococc  
43 320 65.4 395 2 Q9LAY9 streptococc  
44 216 44.2 653 2 Q34097 streptococc  
45 200 40.9 246 2 Q9LSB4 streptococc

#### ALIGNMENTS

RESULT 1  
Q54972 PRELIMINARY; PRT; 619 AA.  
AC Q54972;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Pneumococcal surface protein A precursor.  
GN Name=pspa;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92105030; PubMed=1729249;  
RA Yother J., Briles D.E.;  
RT "Structural properties and evolutionary relationships of PspA, a surface protein of Streptococcus pneumoniae, as revealed by sequence analysis."  
RL J. Bacteriol. 174:601-609(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Yother J., Briles D.E.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
DR EMBL; M74122; AAA27018.1; -;  
DR PIR; A41971; A41971.  
DR PIR; A97887; A97887.  
DR HSSP; P06653; LHCX.  
DR InterPro; IPR002479; CW\_binding.  
DR InterPro; IPR002345; Lipocalin.  
DR Pfam; PF01473; CW\_binding\_1; 10.  
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_3.  
KW Signal.  
FT SIGNAL. 1 31 Potential.  
FT CHAIN 32 639 pneumococcal surface protein A.  
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;  
Query Match 96.0%; Score 469.5; DB 2; Length 619;  
Best Local Similarity 98.0%; Pred. No. 8.6e-21;  
Matches 98; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
QY 1 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQKAA 60  
|||||  
DB 223 LKEIDSESEDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQKAA 281  
QY 61 EENNVEDYFKGLEKTIKAAKAELEKTEADLKKAYPE 100  
|||||  
DB 282 EENNVEDYFKGLEKTIKAAKAELEKTEADLKKAYPE 321  
RESULT 2  
Q8DR10 PRELIMINARY; PRT; 619 AA.

[illegible]

```
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL: AF068647; AAF13457.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW_binding_1; 1.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; Ysirk_signal; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 739 739
SQ SEQUENCE 739 AA; 83960 MW; 7EE2F2P676ABF989 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 739;
Best Local Similarity 86.0%; Pred. No. 1.1e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 537 LKEIDSESDYAKGFRAPLQSKLDTKKAQKLSKLEELSDKIDELDAETAKLEDQLKAA 595
QY 61 EENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 100
DB 596 EGNNVVEAYFKEGLEKTTAAKKAELKTEADLKAVDEPE 635

RESULT 6
Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (fragment).
GN Names:pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL: AF068650; AAF13460.1; -.
DR HSSP: P04268; 1IC2.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW_binding_1; 1.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; Ysirk_signal; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 820 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 820;
Best Local Similarity 86.0%; Pred. No. 1.2e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;
```

```
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 530 LKEIDSESDYAKGFRAPLQSKLDTKKAQKLSKLEELSDKIDELDAETAKLEDQLKAA 588
QY 61 EENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 100
DB 589 EGNNVVEAYFKEGLEKTTAAKKAELKTEADLKAVDEPE 628

RESULT 7
Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=srf10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL: AF154037; AAF73809.1; -.
DR HSSP: P06653; 1H8G.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR005877; Gpos_Ysirk.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW_binding_1; 11.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; Ysirk_signal; 1.
DR TIGRFAMs: TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 929;
Best Local Similarity 86.0%; Pred. No. 1.3e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAETAKLEDQLKAA 60
DB 530 LKEIDSESDYAKGFRAPLQSKLDTKKAQKLSKLEELSDKIDELDAETAKLEDQLKAA 588
QY 61 EENNVEDYFKEGLEKTTAAKKAELKTEADLKAVNEPE 100
DB 589 EGNNVVEAYFKEGLEKTTAAKKAELKTEADLKAVDEPE 628

RESULT 8
Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSPSP; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; YSRK signal; 1.
DR TIGRFAMs; TIGR01168; YSRK signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 82.3%; Score 402.5; DB 2; Length 929;
Best Local Similarity 86.0%; Pred. No. 1.3e-16;
Matches 86; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAA 60
Db 530 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEVLQKD-A 588

Qy 61 EENNVDYFKGLEKTIKAELEKTEADLKKAVNEP 100
Db 589 EGNNVAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 9
Q9LAY4
ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
DR NON TER 437
FT NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 79.7%; Score 389.5; DB 2; Length 437;
Best Local Similarity 84.0%; Pred. No. 4.2e-16;
Matches 84; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAA 60
Db 235 LKEIDSESDYKKEGLRAPLQSKLDTKAKLSKLELSKIDELDAEIAKHVQLQKD-A 293

Qy 61 EENNVDYFKGLEKTIKAELEKTEADLKKAVNEP 100
Db 294 EGNNVAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 333

us-10-674-755-12.rup

```

```

RESULT 10
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
DR NON TER 395
FT NON TER 395
SQ SEQUENCE 395 AA; 42963 MW; 58E6EF956BCBCC1E CRC64;

Query Match 79.2%; Score 387.5; DB 2; Length 395;
Best Local Similarity 83.0%; Pred. No. 5.1e-16;
Matches 83; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAA 60
Db 225 LKEIDSESDYAKGFRAPLQSKLDKAKQAKLSKLELSKIDELDAEIAKLEVLQKD-A 283

Qy 61 EENNVDYFKGLEKTIKAELEKTEADLKKAVNEP 100
Db 284 EGNNVAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 323

RESULT 11
Q9LAY0
ID Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin alc DH like.
DR NON TER 408
FT NON TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 79.2%; Score 387.5; DB 2; Length 408;
Best Local Similarity 83.0%; Pred. No. 5.3e-16;

```

```
Matches 83; Conservative 7; Mismatches 9; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAA 60
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 228 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLEVLQKD-A 286
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 287 EGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 326

RESULT 12
Q8KQK4
ID Q8KQK4 PRELIMINARY; PRT; 99 AA.
AC Q8KQK4;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL: AY082388; AAL92493.1; -.
FT NON_TER 1
FT NON_TER 99
FT NON_TER 99
SQ SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match 78.4%; Score 383.5; DB 2; Length 99;
Best Local Similarity 82.0%; Pred. No. 2.8e-16;
Matches 82; Conservative 5; Mismatches 12; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAA 60
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1 LKEIDSESDYVKEGLRAPLQSELDTKKAKLLKLEELSGKIEELDAEIAEVLQKD-A 59
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 60 EGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVNEPE 99

RESULT 13
Q9L575
ID Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of international
```

```
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 1
FT NON_TER 249
FT NON_TER 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 78.4%; Score 383.5; DB 2; Length 249;
Best Local Similarity 82.0%; Pred. No. 6e-16;
Matches 82; Conservative 6; Mismatches 11; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAA 60
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 74 LKEIDSESDYVKEGLRAPLQSELDTKKAKLLKLEELSGKIEELDAEIAEVLQKD-A 132
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 133 EGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 172

RESULT 14
Q8GNS8
ID Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8;
DT 01-WAR-2003 (TREMBlrel. 23, Created)
DT 01-WAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-WAR-2004 (TREMBlrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=22441996; PubMed=12354862;
RA Dicunzio G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
RT Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
DR HSP; P00192; LAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 224
FT NON_TER 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 77.0%; Score 376.5; DB 2; Length 224;
Best Local Similarity 80.0%; Pred. No. 1.5e-15;
Matches 80; Conservative 8; Mismatches 11; Indels 1; Gaps 1;
QY 1 LKEIDSESDYAKGFRAPLQSKLDAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAA 60
   |:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
Db 17 LKDINSESDYVKEGLRAPLQSELDTKKAKLLKLEELSGKIEELDAEIAEVLQKD-A 75
QY 61 EENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 100
   |:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
Db 76 EGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 115

RESULT 15
Q9LAY5
ID Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBL5;  
RX MEDLINE=20448953; PubMed=10992499;  
RA DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination  
in Streptococcus pneumoniae."  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071810; AAP27706.1; -.  
DR HSSP; P00192; IM6T.  
DR InterPro; IPR011047; Quin\_alc\_DH\_like.  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON TER 426  
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;  
  
Query Match 75.8%; Score 370.5; DB 2; Length 426;  
Best Local Similarity 79.0%; Pred. No. 5.8e-15;  
Matches 79; Conservative 8; Mismatches 12; Indels 1; Gaps 1;  
  
Qy 1 LKEIDSESEDYAKEGFRAPLQSKLDAKQAKLSKLELSKDIDELDAEIAKLEDLKAA 60  
Db 215 LKQINESDSBDYVKEGURAPLQSELDTKKAKLLKLELSGKIEELDAEIAELEVLQKD-A 273  
  
Qy 61 EENNVEDYFKEGLEKTIAAKAELEKTEADLKAVNEPE 100  
Db 274 EGNNVVAYFKEGLEKTIKAELEKAEADLKAVDEPE 313

Search completed: June 21, 2005, 10:22:10  
Job time : 61.9388 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LXEIDSESEDIKSGFRAP.....KKTELEKTEADLKXAVNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 482   | 99.2        | 195    | 2     | AAW14591 Streptoco  |
| 2          | 482   | 99.2        | 195    | 7     | ABW02625 Wu2c pneu  |
| 3          | 482   | 99.2        | 8991   | 6     | ABU08487 S. pneumo  |
| 4          | 453   | 93.2        | 198    | 7     | ABW02615 Rx1c pneu  |
| 5          | 453   | 93.2        | 315    | 2     | AAW04375 Streptoco  |
| 6          | 453   | 93.2        | 619    | 2     | AAW63437 Pneumococ  |
| 7          | 453   | 93.2        | 619    | 2     | AAW87598 Pneumococ  |
| 8          | 453   | 93.2        | 619    | 2     | AAW86911 Pneumococ  |
| 9          | 453   | 93.2        | 619    | 2     | AAW41838 Streptoco  |
| 10         | 453   | 93.2        | 619    | 5     | AAE18782 S. pneumo  |
| 11         | 453   | 93.2        | 619    | 6     | ABU45778 Streptoco  |
| 12         | 453   | 93.2        | 619    | 8     | ADOS2126 Protein e  |
| 13         | 453   | 93.2        | 648    | 2     | AAW70336 Streptoco  |
| 14         | 453   | 93.2        | 648    | 2     | AAW62274 Streptoco  |
| 15         | 453   | 93.2        | 648    | 2     | AAW41837 Streptoco  |
| 16         | 453   | 93.2        | 648    | 2     | AAW87879 A. pneumoc |
| 17         | 453   | 93.2        | 653    | 2     | AAW92456 S. pneumo  |
| 18         | 453   | 93.2        | 684    | 2     | AAW73912 Streptoco  |
| 19         | 450   | 92.6        | 198    | 2     | AAW14581 Streptoco  |
| 20         | 446   | 91.8        | 204    | 2     | AAW14571 Streptoco  |
| 21         | 446   | 91.8        | 204    | 7     | ABW02605 Ef1019c p  |
| 22         | 433   | 89.1        | 653    | 2     | AAW27150 PspA frag  |
| 23         | 430.5 | 88.6        | 289    | 2     | AAW62276 Streptoco  |
| 24         | 430.5 | 88.6        | 289    | 2     | AAW41840 Streptoco  |
| 25         | 430.5 | 88.6        | 289    | 2     | AAW87910 Protein s  |

|    |       |      |      |   |                    |
|----|-------|------|------|---|--------------------|
| 26 | 430.5 | 88.6 | 289  | 2 | AAW92458 S. pneumo |
| 27 | 407   | 83.7 | 623  | 6 | ABU08494 Fragment  |
| 28 | 394   | 81.1 | 1231 | 6 | ABU08490 Fragment  |
| 29 | 385   | 79.2 | 170  | 7 | ABW02614 Rct135c p |
| 30 | 385   | 79.2 | 181  | 7 | ABW02596 0922134c  |
| 31 | 385   | 79.2 | 865  | 6 | ABU08489 S. pneumo |
| 32 | 385   | 79.2 | 929  | 2 | AAW14593 Streptoco |
| 33 | 385   | 79.2 | 929  | 2 | AAW43384 S. pneumo |
| 34 | 382   | 78.6 | 188  | 2 | AAW14580 Streptoco |
| 35 | 382   | 78.6 | 188  | 7 | ABW02613 Rct139c p |
| 36 | 375   | 77.2 | 204  | 2 | AAW14578 Streptoco |
| 37 | 375   | 77.2 | 204  | 7 | ABW02612 Rct123c p |
| 38 | 373   | 76.7 | 588  | 6 | ABU08491 Coiled co |
| 39 | 373   | 76.7 | 588  | 2 | AAW43392 PspC alph |
| 40 | 370.5 | 76.2 | 180  | 2 | AAW14562 Streptoco |
| 41 | 367.5 | 75.6 | 187  | 2 | AAW14579 Streptoco |
| 42 | 358   | 73.7 | 206  | 2 | AAW14574 Streptoco |
| 43 | 358   | 73.7 | 206  | 7 | ABW02608 Db15c pne |
| 44 | 332.5 | 68.4 | 550  | 8 | ADK48356 Streptoco |
| 45 | 332.5 | 68.4 | 550  | 8 | ADR95223 Novel S.  |

## ALIGNMENTS

### RESULT 1

AAW14591

ID AAW14591 standard; protein; 195 AA.

XX AC AAW14591;

XX DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX XX

DE Streptococcus pneumoniae PspA central region.

XX XX

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis; bacteraemia; pneumonia.

XX XX

OS Streptococcus pneumoniae; strain Wu2.

XX XX

PN WO9709994-A1.

XX XX

PD 20-MAR-1997.

XX XX

PF 16-SEP-1996; 96WO-US014819.

XX XX

PR 15-SEP-1995; 95US-00529055.

XX XX

PA (UABR-) UAB RES FOUND.

XX XX

PI Briles DB, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

XX XX

PI Hollingshead S, Tart R, Brooks-Walter A;

XX XX

DR WPI; 1997-202002/18.

XX XX

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used

XX XX

PS in vaccines for protecting animals against S.pneumoniae infection.

XX XX

XX Example 6; Fig 13; 296pp; English.

XX XX

CC This sequence shows the central portion, including the C-terminus of the

XX XX

CC alpha-helix region and some of the proline-rich region, of pneumococcal

XX XX

CC surface protein A (PspA) of Streptococcus pneumoniae strain Wu2.

XX XX

CC Comparison of the N-terminal and central regions (AAW14533-57 and

XX XX

CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can

XX XX

CC be used to divide the strains into several families based on sequence

XX XX

CC homologues. PspA polypeptides, or fragments of them, can be used in

XX XX

CC vaccines to protect animals against S. pneumoniae infection and hence for

XX XX

CC the prevention of diseases such as otitis media, meningitis, bacteraemia

XX XX

CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical

XX XX

CC region and the immediate 5' tip of the coding sequence are likely to be

XX XX

CC the critical sequences for predicting PspA cross-reactions and vaccine

```

CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 195 AA;

Query Match          99.2%; Score 482; DB 2; Length 195;
Best Local Similarity 99.0%; Pred. No. 1.4e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||
Db 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||

Qy 61 ENNVEDYSTGLEKTIAAKKTELEKTEADLKKA VNEPE 99
    |||||
Db 61 ENNVEDYSTGLEKTIAAKKTELEKTEADLKKA VNEPE 99
    |||||

RESULT 2
ABW02625
ID ABW02625 standard; protein; 195 AA.
AC
XX
XX
XX
XX 12-FEB-2004 (first entry)
DT
DE Wu2c pneumococcal surface protein A (PspA) central region.
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
OS Unidentified.
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 71; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Wu2c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
XX Sequence 195 AA;

Query Match          99.2%; Score 482; DB 7; Length 195;
Best Local Similarity 99.0%; Pred. No. 1.4e-35;

CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 195 AA;

Query Match          99.2%; Score 482; DB 2; Length 195;
Best Local Similarity 99.0%; Pred. No. 1.4e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||
Db 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||

Qy 61 ENNVEDYSTGLEKTIAAKKTELEKTEADLKKA VNEPE 99
    |||||
Db 61 ENNVEDYSTGLEKTIAAKKTELEKTEADLKKA VNEPE 99
    |||||

RESULT 2
ABW02625
ID ABW02625 standard; protein; 195 AA.
AC
XX
XX
XX
XX 12-FEB-2004 (first entry)
DT
DE Wu2c pneumococcal surface protein A (PspA) central region.
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
OS Unidentified.
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
XX Example 6; SEQ ID NO 71; 121pp; English.
XX
XX The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Wu2c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
XX Sequence 195 AA;

Query Match          99.2%; Score 482; DB 7; Length 195;
Best Local Similarity 99.0%; Pred. No. 1.4e-35;

Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
    |||||
Db 8797 LKEIDSESEDYAKGFRAPLHSLDKAKOAKLSKLELSKIDELDAEIAKLEDLQKAVE 8856
    |||||

```



QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99  
 DB 8857 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 8895

## RESULT 4

ABW02615  
 ID ABW02615 standard; protein; 198 AA.

XX AC ABW02615;  
 XX DT 12-FEB-2004 (first entry)  
 XX DE RxlC pneumococcal surface protein A (PspA) central region.  
 XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 XX KW immunological; gene therapy; immunostimulant.  
 XX OS Unidentified.

PN US6592876-B1.

PD 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

XX 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 61; 121pp; English.

XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is RxlC pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention

XX SQ Sequence 198 AA;  
 Query Match 93.2%; Score 453; DB 7; Length 198;  
 Best Local Similarity 93.9%; Pred. No. 6.1e-33;  
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEDISESDYAKGFRAPLHSLDKAKLSELSKIDELDAETAKLEDQKAVE 60  
 DB 1 LKEDISESDYAKGFRAPLHSLDKAKLSELSKIDELDAETAKLEDQKAAE 60

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99  
 DB 61 ENNVEDYFKEGLEKTIAAKAELEKTEADLKAVNEPE 99

## RESULT 5

AAV04375  
 ID AAV04375 standard; protein; 315 AA.  
 XX AC AAV04375;  
 XX DT 23-JUN-1999 (first entry)  
 XX DE Streptococcus pneumoniae PspA protein sequence.

XX KW Streptococcus pneumoniae; pspA; pneumococcal; surface protein; vaccine;  
 XX KW immunological; infection.

XX OS Streptococcus pneumoniae.

XX OS Synthetic.

XX PN WO9914333-A2.

XX PD 25-MAR-1999.

XX 18-SEP-1998; 98WO-US019740.

XX 18-SEP-1997; 97US-00932982.

XX (INMR ) PASTEUR MERIEUX CONNAUGHT.

XX PI Becker R, Gray M, Pyle D;

XX WPI; 1999-229537/19.

XX DR N-PSDB; AAX33124.

XX PT DNA encoding PspA molecule with modified internal translational  
 PT initiation sites.  
 XX PS Disclosure; Page; 36pp; English.  
 XX The present sequence represents a pneumococcal surface protein A (PspA)  
 CC molecule where internal naturally occurring translational initiation  
 CC sites have been modified or eliminated so that expression of the DNA  
 CC sequence results in a single form of PspA. The PspA nucleotide sequence  
 CC can be used to transform a unicellular host to produce the PspA protein.  
 CC The PspA protein can be used in an immunological composition for treating  
 CC or preventing S. pneumoniae infection especially in a child. Antibodies  
 CC to the PspA protein can also be used to treat S. pneumoniae infection.  
 CC The immunogenic peptides are designed to confer broad protection against  
 CC diverse pneumococcal strains. N.B. The present sequence is not given in  
 CC the specification but is encoded by the sequence given in AAX33124

XX SQ Sequence 315 AA;

Query Match 93.2%; Score 453; DB 2; Length 315;  
 Best Local Similarity 93.9%; Pred. No. 1.1e-32;  
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEDISESDYAKGFRAPLHSLDKAKLSELSKIDELDAETAKLEDQKAVE 60  
 DB 193 LKEDISESDYAKGFRAPLHSLDKAKLSELSKIDELDAETAKLEDQKAAE 252

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99  
 DB 253 ENNVEDYFKEGLEKTIAAKAELEKTEADLKAVNEPE 291

## RESULT 6

AAV63437  
 ID AAV63437 standard; protein; 619 AA.

XX AC AAV63437;

XX 09-SEP-2004 (revised)

XX 16-OCT-2003 (revised)

XX 25-MAR-2003 (revised)

XX 19-JUL-1995 (first entry)

DE Pneumococcal surface protein A from S.pneumoniae Rx1.  
 XX  
 KW Pneumococcal surface protein A; PspA; Streptococcus; PCR; pneumococcal;  
 KW primer; protection-eliciting epitope; epitope; vaccine; amplify.  
 XX  
 OS Streptococcus pneumoniae.  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 XX Protein 192..260  
 XX /note= "protein fragment of Claim 1"  
 XX  
 XX EP622081-A2.  
 XX  
 XX 02-NOV-1994.  
 XX  
 XX 19-APR-1994; 94EP-00302767.  
 XX  
 XX 20-APR-1993; 93US-00048896.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 XX  
 XX Briles DE, Yother JL, Medaniel LS;  
 XX WPI; 1994-359522/45.  
 XX N-PSDB; AAQ78131.  
 XX  
 XX regions of Pneumococcal surface protein A - derived from the Rx1 PspA  
 XX strain, for the preparation of cross-reactive vaccines for the prevention  
 XX of pneumococcal infections.  
 XX  
 XX Disclosure; Page 13-16; 26pp; English.  
 XX  
 XX The amino acid sequence of the novel Pneumococcal surface protein A  
 XX (PspA) from Streptococcus pneumoniae strain Rx1. The gene was PCR  
 XX amplified from S.pneumoniae genomic DNA using the primers AAQ78132-5. The  
 XX gene was used to derive truncated peptide fragments containing protection  
 XX -eliciting epitopes for use in vaccines against pneumococcal diseases.  
 XX The epitopic fragments are derived from amino acids 192-260 and  
 XX optionally contain a further 25 a.a. residues at both the N- and C-  
 XX terminal regions of the peptide. The epitopic peptide fragments may be  
 XX derived from different strains of S.pneumoniae and are homologous to the  
 XX Rx1 strain epitope. (Updated on 25-MAR-2003 to correct PN field.)  
 XX (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 XX Revised record issued on 09-SEP-2004 : Correction to feature table key  
 XX  
 XX Sequence 619 AA;  
 XX  
 XX Query Match 93.2%; Score 453; DB 2; Length 619;  
 XX Best Local Similarity 93.9%; Pred. No. 2.4e-32;  
 XX Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 LKEIDSESDYAKGFRAPHSKLDKAKQKLSKLELSKIDELDAEIAKLELDQKAVE 60  
 Db 223 LKEIDSESDYAKGFRAPHSKLDKAKQKLSKLELSKIDELDAEIAKLELDQKAAE 282  
 Qy 61 ENNVEDYSTGLEKTTAAKKTLEKTEADLKKA VNEPE 99  
 Db 283 ENNVEDYFREGLEKTTAAKKAELKTEADLKKA VNEPE 321  
 RESULT 7  
 AAR87598  
 ID AAR87598 standard; protein; 619 AA.  
 AC AAR87598;  
 XX  
 XX 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 04-JUL-1996 (first entry)  
 XX  
 DE Pneumococcal surface protein (PspA).

XX PspA; pneumococcal surface protein; truncated; immunoprotective;  
 KW soluble fragment; insertion-duplication mutagenesis.  
 XX  
 OS Streptococcus pneumoniae; strain Rx1.  
 XX  
 XX Key Location/Qualifiers  
 XX Peptide 1..31  
 XX /label= signal\_peptide  
 XX Protein 32..619  
 XX /label= mature\_protein  
 XX Region 32..319  
 XX /label= alpha-helical\_coiled-coil region  
 XX /note= "contains a seven-residue periodicity"  
 XX Region 320..401  
 XX /note= "proline-rich region"  
 XX Region 402..421  
 XX /note= "repeat region"  
 XX Region 422..441  
 XX /note= "repeat region"  
 XX Region 442..461  
 XX /note= "repeat region"  
 XX Region 462..481  
 XX /note= "repeat region"  
 XX Region 482..501  
 XX /note= "repeat region"  
 XX Region 502..521  
 XX /note= "repeat region"  
 XX Region 522..541  
 XX /note= "repeat region"  
 XX Region 542..561  
 XX /note= "repeat region"  
 XX Region 562..581  
 XX /note= "repeat region"  
 XX Region 582..619  
 XX /note= "hydrophobic region starts in last repeat region  
 XX is potential membrane-spanning region"  
 US5476929-A.  
 PD 19-DEC-1995.  
 XX 03-JUN-1993; 93US-00072070.  
 XX 15-FEB-1991; 91US-00656773.  
 XX 12-FEB-1992; 92US-00835698.  
 XX (UABR-) UAB RES FOUND.  
 XX  
 XX Medaniel LS, Yother JL, Briles DE;  
 XX WPI; 1996-049021/05.  
 XX N-PSDB; AAT08979.  
 XX  
 XX New pneumococcal surface protein A fragments - comprise proline-rich  
 XX region and/or repeat region, used for detection of Streptococcus  
 XX pneumoniae.  
 XX  
 XX Claim 1; Col 15-20; 23pp; English.  
 XX  
 XX The present sequence is that of PspA (pneumococcal surface protein A)  
 XX encoded by AAT08979. Through the technique of insertion-duplication  
 XX mutagenesis of the pspA gene of the strain Rx1 of Streptococcus  
 XX pneumoniae with plasmids contg. cloned fragments of the pspA structural  
 XX gene, it has been possible to produce soluble fragments of PspA that are  
 XX secreted by pneumococci. The method can be used to provide an  
 XX immunoprotective truncated PspA protein. Primers and probes based on the  
 XX present sequence are claimed, and are useful for the detection of (at  
 XX least 32) S. pneumoniae strains. (Updated on 25-MAR-2003 to correct PF  
 XX field.) (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 XX Sequence 619 AA;

Query Match 93.2%; Score 453; DB 2; Length 619;  
 Best Local Similarity 93.9%; Pred. No. 2.4e-32;  
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60  
 |||||  
 Db 223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 282  
 |||||

QY 61 ENNVEDYSTEGLEKTIAAKKTLEKTEADLKKAVNEPE 99  
 |||||  
 Db 283 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321  
 |||||

RESULT 8  
 AAR86911  
 ID AAR86911 standard; protein; 619 AA.  
 XX  
 AC AAR86911;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 11-MAY-1996 (first entry)  
 XX  
 XX Pneumococcal surface protein A.  
 DE  
 XX Pneumococcal surface protein A; PspA; cross-protection; vaccine;  
 KW Streptococcus pneumoniae; probe; primer; polymerase chain reaction;  
 KW otitis media; meningitis; bacteraemia; pneumonia; epitope.  
 XX  
 OS Streptococcus pneumoniae; strain Rxl.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..31  
 FT /label= Sig\_peptide  
 FT Region 32..288  
 FT /note= "N-terminal region is highly charged and includes  
 FT an alpha-helix structure"  
 FT Region 289..619  
 FT /note= "C-terminal region includes a proline-rich region  
 FT and a repeat region"  
 XX  
 XX AU9520112-A.  
 XX  
 PD 30-NOV-1995.  
 XX  
 XX 18-MAY-1995; 95AU-00020112.  
 XX  
 PR 20-MAY-1994; 94US-00246636.  
 PR 07-OCT-1994; 94US-00319795.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA Briles DE, Yother JL, McDaniel LS;  
 PI WPI; 1996-030801/04.  
 DR N-PSDB; AAT071178.  
 XX  
 XX Pneumococcal DNA primers and probes - amplify and detect cross-protective  
 PT epitope(s) from Streptococcus pneumoniae surface protein A.  
 XX  
 PS Disclosure; Page 41-43; 61pp; English.  
 XX  
 XX Surface protein A, PspA (AAR86911), of Streptococcus pneumoniae Rxl is  
 CC the product of the pspA gene (AAT07178). PspA includes regions,  
 CC comprising e.g. amino acids 182-588, 293-588 and 192-299, that elicit  
 CC cross-protection against challenge by multiple wild-type strains of S.  
 CC pneumoniae. These cross-reactive epitopes can be prep'd. by expression of  
 CC DNA obtd. by PCR amplification (see AAT07179-96), for use in vaccine  
 CC compns. (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 619 AA;

Query Match 93.2%; Score 453; DB 2; Length 619;  
 Best Local Similarity 93.9%; Pred. No. 2.4e-32;  
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60  
 |||||  
 Db 223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 282  
 |||||

QY 61 ENNVEDYSTEGLEKTIAAKKTLEKTEADLKKAVNEPE 99  
 |||||  
 Db 283 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321  
 |||||

RESULT 9  
 AAY41838  
 ID AAY41838 standard; protein; 619 AA.  
 XX  
 AC AAY41838;  
 XX  
 DT 08-DEC-1999 (first entry)  
 XX  
 DE Streptococcus pneumoniae Rxl PspA protein sequence.  
 XX  
 KW Streptococcus pneumoniae Rxl; PspA; immunoprotective; vaccine; diagnosis;  
 KW infection; pneumococcal surface protein A.  
 XX  
 OS Streptococcus pneumoniae.  
 PN US5965400-A.  
 XX  
 PD 12-OCT-1999.  
 XX  
 PF 23-MAY-1994; 94US-00247491.  
 XX  
 PR 15-FEB-1991; 91US-00656773.  
 PR 12-FEB-1992; 92US-00835698.  
 XX  
 XX (UABR-) UAB RES FOUND.  
 PA Yother JL, Briles DE;  
 PI WPI; 1999-579913/49.  
 DR N-PSDB; AAT25063.  
 XX  
 XX DNA encoding a truncated pneumococcal surface protein A used in the  
 PT development of pneumococcal infections.  
 XX  
 PS Claim 1; Fig 3; 27pp; English.  
 XX  
 CC The present sequence represents Streptococcus pneumoniae Rxl  
 CC immunoprotective Pneumococcal surface protein A (PspA). The present  
 CC invention also describes a method of forming the immunoprotective  
 CC truncated PspA, comprising incorporating a vector comprising the isolated  
 CC DNA molecule encoding PspA (I), into a bacterium via transfection. (I)  
 CC is used to design primers which are capable of detecting a large number  
 CC of S. pneumoniae strains, which in turn can be used to diagnose  
 CC pneumococcal infection in mammals (e.g. humans), independent of the  
 CC strain which has caused it. The PspA protein is used to develop a vaccine  
 CC against pneumococcal infection comprising, as an immunologically-active  
 CC component, a live attenuated or killed bacteria containing a gene coding  
 CC for the truncated form of PspA  
 XX  
 SQ Sequence 619 AA;

Query Match 93.2%; Score 453; DB 2; Length 619;  
 Best Local Similarity 93.9%; Pred. No. 2.4e-32;  
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60  
 |||||  
 Db 223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 282  
 |||||

QY 61 ENNVEDYSTEGLEKTIAAKKTLEKTEADLKKAVNEPE 99  
 |||||  
 Db 283 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321  
 |||||

Query Match 93.2%; Score 453; DB 2; Length 619;  
 Best Local Similarity 93.9%; Pred. No. 2.4e-32;  
 Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

RESULT 10
AAE18782
ID AAE18782 standard; protein; 619 AA.
XX AC
XX AAE18782;
XX DT
XX 17-MAY-2002 (first entry)
XX DE
XX S. pneumoniae Rx1 strain pneumococcal surface protein A (PspA).
XX KW
XX Coiled-coil structural scaffold; heptad repeat; epitope; immune response;
XX cell-mediated immunity; microbial infection; cross-protection; therapy;
XX antimicrobial; vaccine; pneumococcal surface protein A; PspA.
XX OS
XX Streptococcus pneumoniae.
XX FH
XX Key Location/Qualifiers
XX Domain 1..314
XX /label= Helical_domain
XX FT
XX 1..303
XX /note= "N-terminal region"
XX FT
XX 38..44
XX /note= "Immunogenic region 3"
XX FT
XX 40..46
XX /note= "Immunogenic region 5"
XX FT
XX 75..80
XX /note= "Immunogenic region 29"
XX FT
XX 82..87
XX /note= "Immunogenic region 52"
XX FT
XX 96..101
XX /note= "Immunogenic region 66"
XX FT
XX 114..119
XX /note= "Immunogenic region 73"
XX FT
XX 130..135
XX /note= "Immunogenic region 78"
XX FT
XX 137..142
XX /note= "Immunogenic region 89"
XX FT
XX 140..145
XX /note= "Immunogenic region 91"
XX FT
XX 152..156
XX /note= "Immunogenic region 95"
XX FT
XX 153..170
XX /label= Coiled_coil_motif
XX FT
XX 161..164
XX /note= "Immunogenic region 101"
XX FT
XX 166..170
XX /note= "Immunogenic region 116"
XX FT
XX 173..177
XX /note= "Immunogenic region 122"
XX FT
XX 176..180
XX /note= "Immunogenic region 123"
XX FT
XX 181..198
XX /label= Coiled_coil_motif
XX FT
XX 187..191
XX /note= "Immunogenic region 130"
XX FT
XX 194..198
XX /note= "Immunogenic region 133"
XX FT
XX 215..219
XX /note= "Immunogenic region 140"
XX FT
XX 226..230
XX /note= "Immunogenic region 145"
XX FT
XX 229..232
XX /note= "Immunogenic region 148"
XX FT
XX 260..264
XX /note= "Immunogenic region 166"
XX FT
XX 263..267
XX /note= "Immunogenic region 168"
XX FT
XX 284..287
XX /note= "Immunogenic region 76"
XX FT
XX 286..290
XX /note= "Immunogenic region 179"
XX FT

Region 294..298
/Note= "Immunogenic region 182"
FT
Region 295..299
/Note= "Immunogenic region 185"
FT
Region 309..313
/Note= "Immunogenic region 195"
FT
Region 322..326
/Note= "Immunogenic region 206"
FT
XX
XX W0200196368-A2.
XX
XX 20-DEC-2001.
XX
XX 14-JUN-2001; 2001WO-US019168.
XX
XX 14-JUN-2000; 2000US-0211892P.
XX
XX 23-JUN-2000; 2000US-0213387P.
XX
XX (CYTO-) CYTOVAX BIOTECHNOLOGIES INC.
XX
XX Houston ME, Hodges RS;
XX
XX WPI; 2002-188298/24.
XX
XX New synthetic peptide derived from naturally occurring microbial and non-
XX microbial protein antigen useful to stimulate and elicit an immune
XX response in an animal.
XX
XX Example 1; Page 90-92; 99pp; English.
XX
XX The invention relates to the uses of coiled-coil structural scaffold to
XX generate structure-specific peptides, including synthetic peptides
XX derived from naturally occurring microbial and non-microbial protein
XX antigens. The structure of the synthetic peptides utilizes a scaffold of
XX heptad repeat units into which epitopes derived from coiled-coil regions
XX of native proteins are spliced. The resulting peptide has a more stable
XX coiled-coil structure, hence improving presentation of the epitopes in a
XX helical conformation. The peptides of the invention are used to stimulate
XX and elicit an immune response in an animal, as vaccine, to treat or
XX prevent microbial infection by several strains and/or species of
XX microorganism, to provide cross-protection to at least one strain and/or
XX species of microorganism and to stimulate antibody production or cell-
XX mediated immunity to the naturally occurring protein. The present
XX sequence is Streptococcus pneumoniae Rx1 strain pneumococcal surface
XX protein A (PspA) which adopts a coiled-coil structure
XX
XX Sequence 619 AA;
XX
Query Match 93.2%; Score 453; DB 5; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Oy 1 LKEIDSESDYAKGFRAPLHSLKDAKQKLSKLESLDKIDELDAEIAKLDELQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLESLDKIDELDAEIAKLDELQKAAE 282
Oy 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKGLEKTIAAKAELEKTEADLKKAVNEPE 321
XX
RESULT 11
ABU45778
ID ABU45778 standard; protein; 619 AA.
XX AC
XX AEU45778;
XX DT
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by prokaryotic essential gene #31305.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX

```

```
OS Streptococcus pneumoniae.
PN WO200277183-A2.
XX
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 05-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA49648.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 73702; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism's activity; (9) manufacturing an antibiotic; (10) profiling a
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 619 AA;
SQ
Query Match 93.2%; Score 453; DB 6; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 60
DB 223 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 282
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
DB 283 ENNVEDYFKEGLESKTIAAKKALEKTEADLKKAVNEPE 321
RESULT 13
AAW70336
ID AAW70336 standard; protein; 648 AA.
XX
```

```
RESULT 12
ADO52126
ID ADO52126 standard; protein; 619 AA.
XX
XX ADO52126;
AC
XX 12-AUG-2004 (first entry)
DT
XX Streptococcus pneumoniae Rxl PspA protein.
DE
XX Immunogenic composition; vaccine; Th2-type immune response;
KW pneumococcal surface protein A; PspA.
XX
XX Streptococcus pneumoniae.
OS
XX Key Location/Qualifiers
FH Peptide 1..31
FT /label= signal_peptide
FT Protein 32..619
FT /note= "S. pneumoniae Rxl mature PspA protein"
XX
XX US2004101531-A1.
PN
XX 27-MAY-2004.
PD
XX
XX 15-APR-2003; 2003US-00414532.
XX
XX 16-APR-2002; 2002US-0372710P.
PR
XX (CURT/) CURTISS R.
PA (KANG/) KANG H Y.
XX
XX Curtiss R, Kang HY;
XX WPI; 2004-399655/37.
DR N-PSDB; ADO52125.
XX
XX New vaccine comprising a live attenuated strain of pathogenic gram-
PT negative bacteria, useful in eliciting a Th2-type immune response in a
PT vertebrate against pathogens, e.g., helminths, fungi, viruses, protozoans
PT or bacteria.
XX
XX Example 8; SEQ ID NO 72; 94pp; English.
XX
XX The invention relates to immunogenic compositions and vaccines comprising
CC a live attenuated strain of pathogenic gram negative bacteria that
CC secretes an antigen. The vaccine is useful in eliciting a Th2-type immune
CC response in a vertebrate against pathogens, e.g., helminths, fungi,
CC viruses, protozoans or bacteria. The present sequence is Streptococcus
CC pneumoniae Rxl pneumococcal surface protein A (PspA). This sequence is
CC used in the exemplification of the invention.
XX
XX Sequence 619 AA;
SQ
Query Match 93.2%; Score 453; DB 8; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 60
DB 223 LKEIDSESDYAKGFRAPLHSLDKAKQKLSKLEELSDKIDELDAETAKLEDQKAVE 282
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
DB 283 ENNVEDYFKEGLESKTIAAKKALEKTEADLKKAVNEPE 321
RESULT 13
AAW70336
ID AAW70336 standard; protein; 648 AA.
XX
```

```

AC AAW70336;
XX 18-NOV-1998 (first entry)
XX Pneumococcal surface protein A (PspA).
DE Pneumococcal surface protein A gene; PspA; PspA epitope; vaccine;
KW insertion-duplication mutagenesis; otitis media; meningitis; bacteraemia;
XX pneumonia.
XX Streptococcus pneumoniae.
OS Streptococcus pneumoniae.
XX Key Location/Qualifiers
FH Peptide 1..31
FT /note= "Signal peptide"
FT Protein 32..648
FT /note= "PspA"
FT Region 32..319
FT /note= "alpha-helical coil region representing the
FT truncated PspA of the invention "
FT Misc-difference 647
FT /note= "Encoded by AGG"
XX US5804193-A.
XX 08-SEP-1998.
XX 17-MAR-1994; 94US-00214222.
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX (UABR-) UAB RES FOUND.
XX Briles DE, Yother JL;
XX WPI; 1998-505588/43.
XX N-PSDB; AAV33264.
XX Truncated pneumococcal surface protein - useful in vaccines against
XX pneumococcal infection.
XX Example 3; Fig 3A-3C; 22pp; English.
XX The present sequence represents the Streptococcus pneumoniae Rxi
XX pneumococcal surface protein A (PspA). The invention provides a purified
XX truncated form of PspA, formed by an insertion-duplication mutagenesis
XX technique, comprising of the first 288 N-terminal residues of the mature
XX form of wild-type PspA (AAW70336). The truncated PspA contains
XX immunoprotective epitopes of PspA. The invention claims for a vaccine
XX against pneumococcal infection, comprising live-attenuated or killed S.
XX pneumoniae, containing the gene coding for the truncated PspA protein.
XX The truncated protein, optionally conjugated to a poorly immunogenic or
XX nonimmunogenic molecule, is claimed to be useful in vaccines against
XX pneumococcal infection, especially otitis media, meningitis, bacteraemia
XX and pneumonia
XX
XX Sequence 648 AA;
Query Match 93.2%; Score 453; DB 2; Length 648;
Best Local Similarity 93.9%; Pred. No. 2.5e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LKIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKDIDELDAEIAKLEQLKAVE 60
Dy 223 LKIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKDIDELDAEIAKLEQLKAVE 282
Qy 61 ENNVEDYSTGLEKTTAAKTEADLKKAVNEPE 99
Dy 283 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 321
XX
XX RESULT 14
AA41837
ID AAY41837 standard; protein; 648 AA.

```

---

```

AAW62274
ID AAW62274 standard; protein; 648 AA.
XX
XX AAW62274;
XX 22-SEP-1998 (first entry)
XX Streptococcus pneumoniae pspA protein.
DE Streptococcus pneumoniae strain Rxi; pspA; immunoprotective; immunogen;
KW pneumococcal surface protein A; cholera toxin B subunit; fusion protein;
XX antigenic.
XX Streptococcus pneumoniae.
OS Streptococcus pneumoniae.
XX US5753463-A.
XX 19-MAY-1998.
XX 06-JUN-1995; 95US-00469434.
XX 15-FEB-1991; 91US-00656773.
XX 12-FEB-1992; 92US-00835698.
XX 03-JUN-1993; 93US-00072065.
XX (UABR-) UAB RES FOUND.
XX Yother JL, Briles DE;
XX WPI; 1998-311399/27.
XX N-PSDB; AAV39470.
XX Truncated pneumococcal surface protein and cholera toxin B sub-unit
XX fusion protein - useful as an immunogen against Streptococcus pneumoniae.
XX Claim 1; Fig 3; 22pp; English.
XX The present sequence represents the pneumococcal surface protein A (PspA)
XX protein from Streptococcus pneumoniae. A recombinant DNA molecule has
XX been developed which encodes a fusion protein comprising a truncated form
XX of PspA and cholera toxin B subunit (CTB), where the DNA molecule
XX comprises a nucleotide sequence encoding the truncated PspA linked by an
XX in-frame genetic fusion to a CTxB gene, and where the truncated PspA
XX contains immunoprotective epitopes and up to 90% of the whole PspA
XX protein, except for the cell membrane anchor region. The fusion protein
XX is useful for providing an immunogen to protect neonates and children
XX against S.pneumoniae. Most antigenic proteins of this strain are not
XX immunogenic enough to provide protection. The antigenic epitopes of the
XX fusion protein are directed against capsular polysaccharide antigens of
XX S.pneumoniae, specifically it contains the protective epitopes of PspA.
XX The protein can also be used in solid-phase immunoassay assays, since
XX it is readily bound to supports coated with monosialanglioside GM1. The
XX fusion protein is more immunogenic against S.pneumoniae than using PspA
XX alone as the immunogen
XX
XX Sequence 648 AA;
Query Match 93.2%; Score 453; DB 2; Length 648;
Best Local Similarity 93.9%; Pred. No. 2.5e-32;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LKIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKDIDELDAEIAKLEQLKAVE 60
Dy 223 LKIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKDIDELDAEIAKLEQLKAVE 282
Qy 61 ENNVEDYSTGLEKTTAAKTEADLKKAVNEPE 99
Dy 283 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 321
XX
XX RESULT 15
AA41837
ID AAY41837 standard; protein; 648 AA.

```

XX AC AAY41837;  
XX DT 08-DEC-1999 (first entry)  
XX DE Streptococcus pneumoniae Rxl PspA protein sequence.  
XX KW Streptococcus pneumoniae Rxl; PspA; immunoprotective; vaccine; diagnosis;  
XX KW infection; pneumococcal surface protein A.  
XX OS Streptococcus pneumoniae.

Search completed: June 21, 2005, 10:10:14  
Job time : 74.8459 secs

XX PH Key Location/Qualifiers  
XX FT Misc-difference 619. .620  
XX FT /note= "a stop codon is present in the nucleotide  
XX FT sequence at this position"  
XX FT Misc-difference 621. .622  
XX FT /note= "a stop codon is present in the nucleotide  
XX FT sequence at this position"  
XX FT Misc-difference 625. .626  
XX FT /note= "a stop codon is present in the nucleotide  
XX FT sequence at this position"  
XX FT Misc-difference 630. .631  
XX FT /note= "a stop codon is present in the nucleotide  
XX FT sequence at this position"  
XX FT Misc-difference 632. .633  
XX FT /note= "a stop codon is present in the nucleotide  
XX FT sequence at this position"

XX PN US965400-A.  
XX PD 12-OCT-1999.  
XX PF 23-MAY-1994; 94US-00247491.  
XX PR 15-FEB-1991; 91US-00656773.  
XX PR 12-FEB-1992; 92US-00835698.  
XX PA (UABR-) UAB RES FOUND.  
XX PI Yother JL, Briles DE;  
XX WPI; 1999-579913/49.  
XX N-PSDB; AAZ25063.

XX PT DNA encoding a truncated pneumococcal surface protein A used in the  
XX development of pneumococcal infections.  
XX PS Claim 1; Fig 3; 27pp; English.

XX CC The present sequence represents Streptococcus pneumoniae Rxl  
XX immunoprotective Pneumococcal surface protein A (PspA). The present  
XX invention also describes a method of forming the immunoprotective  
XX truncated PspA, comprising incorporating a vector comprising the isolated  
XX DNA molecule encoding PspA (1), into a bacterium via transformation. (1)  
XX is used to design primers which are capable of detecting a large number  
XX of S. pneumoniae strains, which in turn can be used to diagnose  
XX pneumococcal infection in mammals (e.g. humans), independent of the  
XX strain which has caused it. The PspA protein is used to develop a vaccine  
XX against pneumococcal infection comprising, as an immunologically-active  
XX component, a live attenuated or killed bacteria containing a gene coding  
XX for the truncated form of PspA

XX SQ Sequence 648 AA;

Query Match 93.2%; Score 453; DB 2; Length 648;  
Best Local Similarity 93.9%; Pred. No. 2.5e-32;  
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKDIDELDAETAKLEDQLKAVE 60  
DB 223 LKEIDSESDYAKGFRAPLQSKLDAKAKLSKLELSKDIDELDAETAKLEDQLKAAE 282

**inis Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LKEIDSESDYAKGFRAP.....KKTELEKTEADLKKAQNEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents:AA.\*  
1: /cgn2\_6/prodata1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata1/iaa/PCUTUS COMB.pep.\*  
6: /cgn2\_6/prodata1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 486   | 100.0       | 99     | 4  | US-09-147-875A-13 |
| 2          | 482   | 99.2        | 99     | 2  | US-08-710-749-12  |
| 3          | 482   | 99.2        | 195    | 4  | US-08-529-055-71  |
| 4          | 482   | 99.2        | 8991   | 4  | US-08-714-741-32  |
| 5          | 453   | 93.2        | 99     | 2  | US-08-710-749-11  |
| 6          | 453   | 93.2        | 198    | 4  | US-08-529-055-61  |
| 7          | 453   | 93.2        | 619    | 1  | US-08-465-746-2   |
| 8          | 453   | 93.2        | 619    | 1  | US-08-214-164-2   |
| 9          | 453   | 93.2        | 619    | 2  | US-08-467-852A-3  |
| 10         | 453   | 93.2        | 619    | 2  | US-08-246-636-2   |
| 11         | 453   | 93.2        | 619    | 2  | US-08-247-491A-3  |
| 12         | 453   | 93.2        | 619    | 2  | US-08-319-795-2   |
| 13         | 453   | 93.2        | 619    | 2  | US-08-468-985-2   |
| 14         | 453   | 93.2        | 619    | 3  | US-08-312-949-2   |
| 15         | 453   | 93.2        | 648    | 1  | US-08-072-070-2   |
| 16         | 453   | 93.2        | 648    | 1  | US-08-469-434-2   |
| 17         | 453   | 93.2        | 648    | 1  | US-08-214-222-2   |
| 18         | 453   | 93.2        | 648    | 2  | US-08-467-852A-2  |
| 19         | 453   | 93.2        | 648    | 2  | US-08-468-718-2   |
| 20         | 453   | 93.2        | 648    | 2  | US-08-247-491A-2  |
| 21         | 453   | 93.2        | 648    | 3  | US-08-446-201-3   |
| 22         | 453   | 93.2        | 695    | 1  | US-08-127-499A-23 |
| 23         | 453   | 93.2        | 695    | 1  | US-08-482-847-23  |
| 24         | 446.5 | 91.9        | 100    | 4  | US-09-147-875A-12 |
| 25         | 446   | 91.8        | 99     | 2  | US-08-710-749-10  |
| 26         | 446   | 91.8        | 99     | 4  | US-09-147-875A-11 |
| 27         | 446   | 91.8        | 204    | 4  | US-08-529-055-51  |

|    |       |      |      |   |                   |                   |
|----|-------|------|------|---|-------------------|-------------------|
| 28 | 441   | 90.7 | 288  | 3 | US-08-312-949-4   | Sequence 4, Appli |
| 29 | 441   | 90.7 | 288  | 3 | US-08-446-201-4   | Sequence 4, Appli |
| 30 | 430.5 | 88.6 | 289  | 1 | US-08-072-070-4   | Sequence 4, Appli |
| 31 | 430.5 | 88.6 | 289  | 1 | US-08-469-434-4   | Sequence 4, Appli |
| 32 | 430.5 | 88.6 | 289  | 1 | US-08-214-222-4   | Sequence 4, Appli |
| 33 | 430.5 | 88.6 | 289  | 2 | US-08-467-852A-5  | Sequence 5, Appli |
| 34 | 430.5 | 88.6 | 289  | 2 | US-08-468-718-4   | Sequence 4, Appli |
| 35 | 430.5 | 88.6 | 289  | 2 | US-08-247-491A-5  | Sequence 5, Appli |
| 36 | 407   | 83.7 | 623  | 4 | US-08-714-741-47  | Sequence 17, Appl |
| 37 | 399   | 82.1 | 99   | 2 | US-08-710-749-17  | Sequence 41, Appl |
| 38 | 394   | 81.1 | 1231 | 4 | US-09-147-875A-10 | Sequence 10, Appl |
| 39 | 389.5 | 80.1 | 100  | 4 | US-09-147-875A-10 | Sequence 10, Appl |
| 40 | 388   | 79.8 | 99   | 2 | US-08-710-749-15  | Sequence 15, Appl |
| 41 | 385   | 79.2 | 170  | 4 | US-08-529-055-60  | Sequence 60, Appl |
| 42 | 385   | 79.2 | 181  | 4 | US-08-529-055-42  | Sequence 42, Appl |
| 43 | 385   | 79.2 | 864  | 4 | US-08-714-741-40  | Sequence 40, Appl |
| 44 | 382   | 78.6 | 99   | 4 | US-09-147-875A-16 | Sequence 16, Appl |
| 45 | 382   | 78.6 | 188  | 4 | US-08-529-055-59  | Sequence 59, Appl |

ALIGNMENTS

RESULT 1  
US-09-147-875A-13  
; Sequence 13, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-13

Query Match : 100.0%; Score 486; DB 4; Length 99;  
Best Local Similarity 100.0%; Pred. No. 1.9e-37;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LKEIDSESDYAKGFRAPLHSLDKAKLSKLEELSDKIDELDAETAKLEDQLKAVE 60  
Db 1 LKEIDSESDYAKGFRAPLHSLDKAKLSKLEELSDKIDELDAETAKLEDQLKAVE 60  
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAQNEPE 99  
Db 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAQNEPE 99

RESULT 2  
US-08-710-749-12  
; Sequence 12, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-12

Query Match 99.2%; Score 482; DB 2; Length 99;  
Best Local Similarity 99.0%; Pred. No. 4.3e-37;  
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAIAKLEDLQKAVE 60  
Db 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAIAKLEDLQKAVE 60  
Qy 61 ENNVEDYSTGLEKTTAAKTELEKTEADLKKA VNEPE 99  
Db 61 ENNVEDYSTGLEKTTAAKTELEKTEADLKKA VNEPE 99

## RESULT 3

US-08-529-055-71  
Sequence 71, Application US/08529055  
Patent No. 6592876  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: Pneumococcal Genes, Portions  
TITLE OF INVENTION: Thereof, Expression Products  
TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
TITLE OF INVENTION: Portions and Products  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 71:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 195 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-71

Query Match 99.2%; Score 482; DB 4; Length 195;  
Best Local Similarity 99.0%; Pred. No. 9.8e-37;  
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAIAKLEDLQKAVE 60  
Db 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSDKIDELDAIAKLEDLQKAVE 60  
Qy 61 ENNVEDYSTGLEKTTAAKTELEKTEADLKKA VNEPE 99  
Db 61 ENNVEDYSTGLEKTTAAKTELEKTEADLKKA VNEPE 99

## RESULT 4

US-08-714-741-32  
Sequence 32, Application US/08714741  
Patent No. 6500613  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J.  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
TITLE OF INVENTION: PORTIONS AND PRODUCTS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/714,741  
FILING DATE: 16-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer Esq., William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2460  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8991 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-714-741-32

```
Query Match          99.2%; Score 482; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 9.8e-35;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 8797 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 8856
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 8857 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 8895

RESULT 5
US-08-710-749-11
; Sequence 11, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-11

Query Match          93.2%; Score 453; DB 2; Length 99;
Best Local Similarity 93.9%; Pred. No. 1.9e-34;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 60
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 61 ENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEPE 99

RESULT 6
US-08-529-055-61
; Sequence 61, Application US/08529055
; Patent No. 6592876

Query Match          93.2%; Score 453; DB 2; Length 99;
Best Local Similarity 93.9%; Pred. No. 1.9e-34;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 60
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 61 ENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEPE 99

RESULT 7
US-08-465-746-2
; Sequence 2, Application US/08465746
; Patent No. 5679768
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: Yother, Janet S.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
```

```
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-61

Query Match          93.2%; Score 453; DB 4; Length 198;
Best Local Similarity 93.9%; Pred. No. 4.4e-34;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAVE 60
    |||||
Db 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDQLKAAE 60
    |||||
QY 61 ENNVEDYSTEGLEKTIAAKTEADLKKAVNEPE 99
    |||||
Db 61 ENNVEDYFKGLEKTIAAKKALEKTEADLKKAVNEPE 99

RESULT 7
US-08-465-746-2
; Sequence 2, Application US/08465746
; Patent No. 5679768
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Yother, Janet L.
; APPLICANT: Yother, Janet S.
; APPLICANT: McDaniel, Larry S.
; TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEIN A
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
```

```
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,746
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/048,896
; FILING DATE:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-746-2

Query Match          93.2%; Score 453; DB 1; Length 619;
Best Local Similarity 93.9%; Pred. No. 1.7e-33;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSHKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSKIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTGLEKTTIAKKTELEKTEADLKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTTIAKKAELKTEADLKAVNEPE 321

RESULT 9
US-08-467-852A-3
; Sequence 3, Application US/08467852A
; Patent No. 5856170
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: YOTHER, Janet L.
; APPLICANT: MCDANIEL, Larry S.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,852A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-467-852A-3

Query Match          93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 1.7e-33;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,746
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/048,896
; FILING DATE:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 619 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-746-2

Query Match          93.2%; Score 453; DB 1; Length 619;
Best Local Similarity 93.9%; Pred. No. 1.7e-33;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSHKLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKAKLSKLELSKIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTGLEKTTIAKKTELEKTEADLKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTTIAKKAELKTEADLKAVNEPE 321

RESULT 8
US-08-214-164-2
; Sequence 2, Application US/08214164
; Patent No. 5728387
; GENERAL INFORMATION:
; APPLICANT: BRILES, DAVID E.
; APPLICANT: YOTHER, JANET L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,164
; FILING DATE: 17-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
```

|    |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |   |   |   |   |   |   |   |    |
|----|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|---|---|---|---|---|---|---|----|
| Qy | 1   | L | K | E | I | D | E | S | E | S | E | D | Y | A | K | E | G | F | R | A | P | L | H | S | K | L | D | A | Q | A | K | S | K | L | E   | L | S | D | K | I | D | E | L | D | A | E | I | A | K | L | E   | D | O | L | K | A | N | E | 60 |
| Db | 223 | L | K | E | I | D | E | S | E | S | E | D | Y | A | K | E | G | F | R | A | P | L | H | S | K | L | E | L | S | D | K | I | D | E | L   | D | A | E | I | A | K | L | E | D | O | L | K | A | N | E | 282 |   |   |   |   |   |   |   |    |
| Qy | 61  | E | N | N | V | D | Y | T | E | G | L | E | K | T | I | A | A | K | T | E | L | E | K | T | E | A | D | L | K | A | N | E | P | E | 99  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |   |   |   |   |   |   |   |    |
| Db | 283 | E | N | N | V | D | Y | T | E | G | L | E | K | T | I | A | A | K | T | E | L | E | K | T | E | A | D | L | K | A | N | E | P | E | 321 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |   |   |   |   |   |   |   |    |

RESULT 10  
 US-08-246-2  
 Sequence 2, Application US/08246636  
 Patent No. 5965141  
 GENERAL INFORMATION:  
 APPLICANT: Briles, David E  
 APPLICANT: Yother, Janet L  
 APPLICANT: McDaniel, Larry S  
 APPLICANT: Wu, Hong-Yin  
 TITLE OF INVENTION: EPITOPIC REGIONS OF PNEUMOCOCCAL SURFACE  
 TITLE OF INVENTION: PROTEIN A  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 :

CONTRACTOR: SHOENAKER AND MATTARE, LTD  
ADDRESS: SUITE 1203, 2001 JEFFERSON DAVIS HIGHWAY  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22202-0286

? 42202 0200  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/08/246,636  
 ? FILING DATE: 20-MAY-1994  
 ?

CLASSIFICATION: 435  
 PRIOR APPLICATION DATA: US 07/656,773  
 FILING DATE: 15-FEB-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/835,698  
 FILING DATE: 12-FEB-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/048,896  
 FILING DATE: 20-APR-1993  
 TELECOMMUNICATION INFORMATION:

TELECOMMUNICATION INFORMATION  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813  
TELEX: LUKPAT WASHINGTON  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 619 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

|                       |       |              |          |            |    |        |      |
|-----------------------|-------|--------------|----------|------------|----|--------|------|
| Query Match           | 93.2% | Score        | 453;     | DB         | 2; | Length | 619; |
| Best Local Similarity | 93.9% | Pred. No.    | 1.7e-33; |            |    |        |      |
| Matches               | 93;   | Conservative | 1;       | Mismatches | 5; | Indels | 0;   |
|                       |       |              |          |            |    | Gaps   | 0;   |

|    |  |     |   |     |
|----|--|-----|---|-----|
| Qy |  | 1   | LKEIDSESEDYAKEGFRAPLHSHKLDKAKQAKLSKLEELSDKIDELDAEIAKLEDQLKAVE | 60  |
|    |  |     | :   |     |
| D6 |  | 223 | LKEIDSESEDYAKEGFRAPLQSKLDKAKQAKLSKLEELSDKIDELDAEIAKLEDQLKAAE  | 282 |

|                      |  |
|----------------------|--|
| <b>Qy</b>            | 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKAVNEPE 99<br> |
| <b>D<sub>b</sub></b> | 283 ENNVEDYFKEGLEKTTAAKAELEKTEADLKAVNEPE 321   |

## RESULT 11

US-08-247-491A-3  
; Sequence 3, Application US/08247491A  
; Patent No. 5965400  
; GENERAL INFORMATION:  
; APPLICANT: BRILES, David E.  
; APPLICANT: YOTHER, Janet L.  
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
; STREET: 745 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA

```

?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/247,491A
? FILING DATE: 23-JUN-1994
?
?
```

FILED DATE: 25-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KOWALSKI, Thomas J.  
REGISTRATION NUMBER: 32,147  
REFERENCE/DOCKET NUMBER: 454312-2041  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-588-0800  
TELEFAX: 212-588-0500

```

; ELEMFA: 242-368-0500
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 619 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: n/a
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: amino acid
US-08-247-491A-3

```

|                       |              |                    |               |             |
|-----------------------|--------------|--------------------|---------------|-------------|
| Query Match           | 93.2%        | Score 453;         | DB 2;         | Length 619; |
| Best Local Similarity | 93.9%        | Pred. No. 1.7e-33; |               |             |
| Matches 93;           | Conservative | 1;                 | Mismatches 5; | Indels 0;   |
| Gaps                  |              |                    |               |             |

|    |     |  |     |
|----|-----|--|-----|
| Qy | 1   | LKETDESESDYAKEGFRAPLHSLDKAAOKLSKLEELSDKTDDEDAETAKLEDOUKAVE | 60  |
| Db | 223 | LKEIDSESDYAKEGFRAPLQSKUDAKAKLSKUEELSDKTDDEDAETAKLEDOUKAAE  | 282 |
| Qy | 61  | ENNNVEDYSTGLEKTTIAAKKTELEKTRADLUKAVNEPE                    | 99  |
| Db | 283 | ENNNVEDYFKGLEKTTAAKKAELKTRADLUKAVNEPE                      | 321 |

RESULT 12  
US-08-319-795-2  
; Sequence 2, Application US/08319795  
; Patent No. 5980909

```

1  GENERAL INFORMATION:
2  APPLICANT: Briles, David E.
3  APPLICANT: Yother, Janet L.
4  APPLICANT: McDaniel, Larry S
5  TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface
6  TITLE OF INVENTION: Protein A
7  NUMBER OF SEQUENCES: 20
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: Sheomaker and Mattare, Ltd.
10 STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson
11 STREET: Davis Highway
12 CITY: Arlington
13 STATE: Virginia
14 COUNTRY: U.S.A.
15 ZIP: 22202-0286
16 COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/319,795  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER: US 08/246,636  
FILING DATE: 20-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/048,896  
FILING DATE: 20-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/835,698  
FILING DATE: 12-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/656,773  
FILING DATE: 15-FEB-1991  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 619 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-319-795-2

Query Match 93.2%; Score 453; DB 2; Length 619;  
Best Local Similarity 93.9%; Pred. No. 1.7e-33;  
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSKIDELDAEIAKLEDLKAVE 60  
Db |||||  
223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSKIDELDAEIAKLEDLKAVE 282

Qy 61 ENNVEDYSTGEGLEKTTAAKTEADLKKAVNEPE 99  
Db |||||  
283 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 321

RESULT 13  
US-08-468-985-2  
Sequence 2, Application US/08468985  
Patent No. 5997882  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: Yother, Janet L.  
APPLICANT: McDaniel, Larry S  
TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheomaker and Mattare, Ltd.  
STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,985  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/319,795  
FILING DATE:  
APPLICATION NUMBER: US 08/246,636  
FILING DATE: 20-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/048,896  
FILING DATE: 20-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/835,698  
FILING DATE: 12-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/656,773  
FILING DATE: 15-FEB-1991  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 619 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-468-985-2

Query Match 93.2%; Score 453; DB 2; Length 619;  
Best Local Similarity 93.9%; Pred. No. 1.7e-33;  
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSKIDELDAEIAKLEDLKAVE 60  
Db |||||  
223 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSKIDELDAEIAKLEDLKAVE 282

Qy 61 ENNVEDYSTGEGLEKTTAAKTEADLKKAVNEPE 99  
Db |||||  
283 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 321

RESULT 14  
US-08-312-949-2  
Sequence 2, Application US/08312949  
Patent No. 6027734  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: Wu, Hong-Yin  
TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF  
TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/312,949  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 619 amino acids

; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-312-949-2

Query Match 93.2%; Score 453, DB 3; Length 619;  
Best Local Similarity 93.9%; Pred. No. 1.7e-33;  
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDESESEDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDDQLKAVE 60  
|||||  
DB 223 LKIDESESEDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDDQLKAAE 282  
|||||  
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99  
|||||  
DB 283 ENNVEDYFKEGLEKTIAAKKALEKTEADLKKAVNEPE 321  
|||||

## RESULT 15

US-08-072-070-2  
; Sequence 2, Application US/08072070  
; Patent No. 5476929  
; GENERAL INFORMATION:  
; APPLICANT: Biles, David E  
; APPLICANT: Yocher, Janet L  
; APPLICANT: McDaniel, Larry S  
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd  
; STREET: Suite 1203, 2001 Jefferson Davis Highway  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/072,070  
; FILING DATE: 19930603  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/835,698  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/656,773  
; FILING DATE: 15-FEB-1991  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 415-0810  
; TELEFAX: (703) 521-0378  
; TELEX: LUKPAT WASHINGTON  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 648 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-072-070-2

Query Match 93.2%; Score 453, DB 1; Length 648;  
Best Local Similarity 93.9%; Pred. No. 1.8e-33;  
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDESESEDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDDQLKAVE 60  
|||||  
DB 223 LKIDESESEDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDDQLKAAE 282  
|||||  
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99  
|||||

DB 283 ENNVEDYFKEGLEKTIAAKKALEKTEADLKKAVNEPE 321  
|||||

Search completed: June 21, 2005, 10:25:20  
Job time : 19.4867 secs

**This Page Blank (uspto)**



## RESULT 2

```
US-10-299-636-86
; Sequence 86, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-86

Query Match      99.2%; Score 482; DB 15; Length 195;
Best Local Similarity 99.0%; Pred. No. 3.1e-31;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db      1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60

Qy      61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99
Db      61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99

RESULT 3
US-10-299-636-76
; Sequence 76, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-76

Query Match      93.2%; Score 453; DB 15; Length 198;
Best Local Similarity 93.9%; Pred. No. 6.8e-29;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db      1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60

Qy      61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99
Db      61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99

RESULT 4
US-10-299-636-105
; Sequence 105, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-105

Query Match      93.2%; Score 453; DB 15; Length 354;
Best Local Similarity 93.9%; Pred. No. 1.3e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db      175 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLEDLQKAAE 234

Qy      61 ENNVEDYSTEGLEKTIAAKKTEADLKKA VNEPE 99
Db      235 ENNVEDYFKEGLEKTIAAKKA E LKTEADLKKA VNEPE 273

RESULT 5
US-10-299-636-96
; Sequence 96, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
```

```
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-96

Query Match          93.2%; Score 453; DB 15; Length 588;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAVE 60
Db 192 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAAE 251

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 252 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 290

RESULT 6
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; TITLE OF INVENTION: Structure-Specific Peptides
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

Query Match          93.2%; Score 453; DB 10; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAAE 282

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 7
US-10-282-122A-73702
; Sequence 73702, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

Query Match          93.2%; Score 453; DB 16; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73702
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73702

Query Match          93.2%; Score 453; DB 15; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLEELSDKIDELDAEIAKLEDLQKAAE 282

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKKAELKTEADLKKAVNEPE 321

RESULT 8
US-10-414-532-72
; Sequence 72, Application US/10414532
; Publication No. US20040101531A1
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; TITLE OF INVENTION: CARRIER BACTERIA THAT SECRETE ANTIGENS
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 72
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-414-532-72

Query Match          93.2%; Score 453; DB 16; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.4e-28;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 223 LKEIDSESDYAKGFRAPLQSLDKAKAKLSKLELSKIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 321

RESULT 9
US-10-674-755-12
; Sequence 12, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-12

Query Match 91.9%; Score 446.5; DB 15; Length 100;
Best Local Similarity 94.0%; Pred. No. 1.1e-28;
Matches 94; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAV 59
Db 1 LKEIDSESDYAKGFRAPLQSLDKAKAKLSKLELSKIDELDAEIAKLEDLQKAA 60

Qy 60 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 100

RESULT 10
US-10-674-755-11
; Sequence 11, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-11

Query Match 91.8%; Score 446; DB 15; Length 99;
Best Local Similarity 91.9%; Pred. No. 1.1e-28;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 1 LKEIDSESDYKVEGFRAPLQSLDKAKAKLSKLELSKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 99
```

```
Db 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 99

RESULT 11
US-10-299-636-66
; Sequence 66, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Iotcher, Janec
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-66

Query Match 91.8%; Score 446; DB 15; Length 204;
Best Local Similarity 91.9%; Pred. No. 2.6e-28;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAVE 60
Db 1 LKEIDSESDYKVEGFRAPLQSLDKAKAKLSKLELSKIDELDAEIAKLEDLQKAAE 60

Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 61 ENNVEDYFKEGLEKTTAAKAELEKTEADLKKAVNEPE 99

RESULT 12
US-10-674-755-10
; Sequence 10, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-10

Query Match 80.1%; Score 389.5; DB 15; Length 100;
Best Local Similarity 84.0%; Pred. No. 4.1e-24;
Matches 84; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKIDELDAEIAKLEDLQKAV 59
Db 1 LKEIDSESDYKVEGFRAPLQSLDKAKAKLSKLELSKIDELDAEIAKLEDLQKAAE 60
```

```
QY 60 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
DB 61 EGNVVEAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 100

RESULT 13
US-10-299-636-75
; Sequence 75, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-75

Query Match 79.2%; Score 385; DB 15; Length 170;
Best Local Similarity 81.8%; Pred. No. 1.7e-23;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLEELSDKIDELDAEIAKLEVDQLKAVE 60
DB 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLEELSDKIDELDAEIAKLEVDQLKAVE 60

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
DB 61 EGNVVEAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 99

RESULT 14
US-10-299-636-57
; Sequence 57, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299, 636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-57

Query Match 79.2%; Score 385; DB 15; Length 181;
Best Local Similarity 81.8%; Pred. No. 1.8e-23;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLEELSDKIDELDAEIAKLEVDQLKAVE 60
DB 1 LKEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLEELSDKIDELDAEIAKLEVDQLKAVE 60

QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKAVNEPE 99
DB 61 EGNVVEAYFKEGLEKTTAEKKAELKAEADLKAVDEPE 99

Search completed: June 21, 2005, 11:18:34
Job time : 63.2388 sec
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LKEIDSESEDYAKGFRAP.....KKTELEKTEADLKAVNEPE 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 453   | 93.2        | 619    | 2 A97887 | surface protein ps  |
| 2          | 453   | 93.2        | 619    | 2 A41971 | surface protein ps  |
| 3          | 129.5 | 26.6        | 744    | 2 F95013 | pneumococcal surfa  |
| 4          | 111.5 | 22.9        | 3488   | 2 T34418 | hypothetical prote  |
| 5          | 110   | 22.6        | 281    | 2 F75216 | hypothetical prote  |
| 6          | 109.5 | 22.5        | 852    | 2 D72330 | conserved hypothet  |
| 7          | 108   | 22.2        | 886    | 2 H69378 | conserved hypothet  |
| 8          | 107.5 | 22.1        | 161    | 2 S48396 | tropomyosin TPM2 -  |
| 9          | 106.5 | 21.9        | 388    | 2 S52536 | fcrA 15 protein -   |
| 10         | 106.5 | 21.9        | 405    | 2 A33339 | Fc gamma (15G) rec  |
| 11         | 105   | 21.6        | 229    | 2 S70532 | outer surface prot  |
| 12         | 105   | 21.6        | 764    | 2 T05049 | hypothetical prote  |
| 13         | 104.5 | 21.5        | 372    | 2 S23326 | gene ML2.2 protein  |
| 14         | 104.5 | 21.5        | 415    | 2 S35760 | fcrA protein precu  |
| 15         | 104.5 | 21.5        | 501    | 2 A44643 | M protein precu     |
| 16         | 103.5 | 21.3        | 387    | 2 S57834 | fcrA protein precu  |
| 17         | 102.5 | 21.1        | 388    | 2 A46173 | Mrp4 protein - Str  |
| 18         | 101   | 20.8        | 402    | 2 S37046 | IgA receptor - Str  |
| 19         | 100.5 | 20.7        | 1312   | 2 T30845 | probable DNA repla  |
| 20         | 100   | 20.6        | 166    | 2 S73342 | hypothetical prote  |
| 21         | 100   | 20.6        | 365    | 2 B54128 | Fc-binding protein  |
| 22         | 100   | 20.6        | 376    | 2 A43528 | protein H precursor |
| 23         | 100   | 20.6        | 1957   | 2 T38077 | hypothetical coile  |
| 24         | 99.5  | 20.5        | 1277   | 2 S53043 | probable membrane   |
| 25         | 98.5  | 20.3        | 583    | 1 A46127 | radixin - human     |
| 26         | 98.5  | 20.3        | 583    | 1 S39805 | radixin - pig       |
| 27         | 98.5  | 20.3        | 1790   | 2 S67593 | transport protein   |
| 28         | 98.5  | 20.3        | 2401   | 2 T28676 | thopry protein -    |
| 29         | 98    | 20.2        | 473    | 2 F70031 | cell wall-binding   |

|    |      |      |      |          |                     |
|----|------|------|------|----------|---------------------|
| 30 | 98   | 20.2 | 1875 | 2 S38173 | myosin-like protei  |
| 31 | 98   | 20.2 | 1937 | 2 I38055 | myosin heavy chain  |
| 32 | 98   | 20.2 | 1957 | 2 A59294 | skeletal myosin -   |
| 33 | 98   | 20.2 | 1979 | 2 C71622 | hypothetical prote  |
| 34 | 97.5 | 20.1 | 384  | 2 S49550 | M-like protein emm  |
| 35 | 97.5 | 20.1 | 880  | 2 F75103 | conserved hypothet  |
| 36 | 97.5 | 20.1 | 896  | 2 S43074 | epidermal growth f  |
| 37 | 97   | 20.0 | 1039 | 2 S62509 | probable vesicular  |
| 38 | 97   | 20.0 | 1044 | 2 T50213 | probable vesicular  |
| 39 | 97   | 20.0 | 1156 | 2 B70356 | chromosome assembl  |
| 40 | 97   | 20.0 | 1177 | 2 B75150 | chromosome segrega  |
| 41 | 97   | 20.0 | 1319 | 2 A28313 | glued protein - fr  |
| 42 | 97   | 20.0 | 1940 | 1 S04090 | myosin heavy chain  |
| 43 | 96.5 | 19.9 | 230  | 2 I40287 | outer surface prot  |
| 44 | 96.5 | 19.9 | 408  | 2 S30283 | protein M precursor |
| 45 | 96.5 | 19.9 | 713  | 2 B84583 | hypothetical prote  |

ALIGNMENTS

RESULT 1

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Mc  
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5703-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: A97887

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <KUR>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl

C:Genetics:

A:Gene: pspA

Query Match 93.2%; Score 453; DB 2; Length 619;  
Best Local Similarity 93.9%; Pred. No. 9.1e-24;  
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDSESEDYAKGFRAPLHSLDKAKLSKLEELSDKIDELDAETAKLEDQLKAVE 60

Db 223 LKEIDSESEDYAKGFRAPLQSKLDKAKLSKLEELSDKIDELDAETAKLEDQLKAAE 282

QY 61 ENNVEDYSTEGLKTTAAKTELEKTEADLKAVNEPE 99

Db 283 ENNVEDYFKEGLKTTAAKAELEKTEADLKAVNEPE 321

RESULT 2

A41971

surface protein pspA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134

R:Yother, J.; Briles, D.E.

J. Bacteriol. 174, 601-609, 1992

A:Title: Structural properties and evolutionary relationships of PspA, a surface protein

A:Reference number: A41971; MUID:92105030; PMID:1729249

A:Accession: A41971

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <YOT>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:g153840; PIDN:AAA2701f

A>Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIIP:75636)

R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Infect. Immun. 59, 1285-1289, 1991  
A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability  
A;Reference number: A60282; MUID:91169598; PMID:2004810  
A;Accession: A60282  
A;Molecule type: protein  
A;Residues: 32-76 <TAL>  
A;Experimental source: strain JY2008  
C;Genetics:  
A;Gene: pspA  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-619/Product: surface protein pspA #status predicted <MAT>  
F:411-430/Domain: cpl repeat homology <CP01>  
F:431-450/Domain: cpl repeat homology <CP02>  
F:451-470/Domain: cpl repeat homology <CP03>  
F:471-490/Domain: cpl repeat homology <CP04>  
F:491-510/Domain: cpl repeat homology <CP05>  
F:511-530/Domain: cpl repeat homology <CP06>  
F:531-550/Domain: cpl repeat homology <CP07>  
F:551-570/Domain: cpl repeat homology <CP08>  
F:571-591/Domain: cpl repeat homology <CP09>  
F:592-611/Domain: cpl repeat homology <CP10>  
Query Match 93.2%; Score 453; DB 2; Length 619;  
Best Local Similarity 93.9%; Pred. No. 9.1e-24;  
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LKEIDSESEDYAKGFRAPLHSLDKDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAVE 60  
Db 223 LKEIDSESEDYAKGFRAPLQSLQSLDKDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAAE 282  
Qy 61 ENNVEDYSTGLEKTIKAAKTELEKTEADLKKAVNEPE 99  
Db 283 ENNVEDYFKEGLSKTIKAAKAELEKTEADLKKAVNEPE 321  
RESULT 3  
F95013  
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: F95013  
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap  
ple, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: F95013  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-744 <KUR>  
A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:G  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0117  
Query Match 26.6%; Score 129.5; DB 2; Length 744;  
Best Local Similarity 36.1%; Pred. No. 0.11;  
Matches 39; Conservative 12; Mismatches 28; Indels 29; Gaps 3;  
Qy 19 APLHSLDKDAKQAKLSK-----LEELSDKI-----DELDAEIAKLEDLQKAVEENNVED 67  
Db 333 AALQNLAAKKAELAKQTELEKLLDPSGKTQDELKREAEAEALDKKADELQNKVAD 392  
Qy 68 Y-----STEGLEKTIKAAKTELEKTEADLKKAVNE 97  
Db 393 LEKEISNLEILLGADSEDDTAALQNLKATKKAELEKTKQKELDAALNE 440  
RESULT 4  
T34418  
hypothetical protein F12F3.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C;Accession: T34418  
R;Fulton, B.; Wohldmann, P.  
Submitted to the EMBL Data Library, July 1998  
A;Description: The sequence of C. elegans cosmid F12F3.  
A;Reference number: Z21521  
A;Accession: T34418  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-3488 <FUL>  
A;Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3  
A;Experimental source: strain Bristol N2; clone F12F3  
C;Genetics:  
A;Gene: CESP:F12F3.3  
A;Map position: 5  
A;Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1  
Query Match 22.9%; Score 111.5; DB 2; Length 3488;  
Best Local Similarity 35.2%; Pred. No. 8.4;  
Matches 44; Conservative 17; Mismatches 29; Indels 35; Gaps 7;  
Qy 2 KEIDES----ESEDYAKGFRAPLHSLDKDAKQAKLSK-----EELSDKIDELDAE 49  
Db 1009 KETDEKILKLDAAEIAAKTKQEADEKSKLDA-QEKIKKVSDDAARKEKELNDKL-KLSEI 1066  
Qy 50 A-----KLQDLKA-----VEENNVEDYSTEGLEKTIKAAKTELEKTEA 89  
Db 1067 ATKASADKULKEBQAQAKAAVEAAKQKQEKDEQLKLDTEAAKAAAKLELEK-QA 1125  
Qy 90 DLKKA 94  
Db 1126 QIKKA 1130  
RESULT 5  
F75216  
hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: F75216  
R;anonymous, Genoscope  
Submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A;Reference number: A75001  
A;Accession: F75216  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-281 <KAW>  
A;Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:95457433; PIDN:CAB49181  
A;Experimental source: strain Orsay  
C;Genetics:  
A;Gene: PAB2181  
Query Match 22.6%; Score 110; DB 2; Length 281;  
Best Local Similarity 27.9%; Pred. No. 0.85;  
Matches 29; Conservative 26; Mismatches 39; Indels 10; Gaps 2;  
Qy 2 KEIDSESEDYAKGFRAPLHSLDKDAKQAKLSKLELSKDIDELDAEIAKLEDLQKAVEE 61  
Db 163 KEIEELKGVEKLEQEKLEKLEKSEVKLMYEAKAKAAEELAKLREYEEKVRE 222  
Qy 62 -----NNNVEDYSTEGLEKTIKAAKTELEKTEADLKKAVNE 97  
Db 223 LERKVELERSLNEYTK--VKSLEKKCELEKNKVELEEVNK 264  
RESULT 6  
D72230  
conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
C;Species: Thermotoga maritima  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C;Accession: D72230



R.; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Thermoplasma acidophilum*

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: D72230

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-852 <ARN>

A;Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982156; PIDN:AAD3670

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: Tm636

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.5%; Score 109.5; DB 2; Length 852;

Best Local Similarity 29.2%; Pred. No. 2.8;

Matches 26; Conservative 22; Mismatches 30; Indels 11; Gaps 2;

QY 6 ESESDYAKGFRAPLHSLKDAKQAKLSKLE---LSDKIDELDAEIAKLEQDLKXAVEE 61

DB 506 EKIEELHRLGYSEDLEKLDKRRKLRKIEERHSISQKITAADVQISQENQLKEIKG 565

QY 62 NNNVEDYSTGLEKTTAAKTELEKTEAD 90

DB 566 -----EIEAKRETLKEORENDOLKSD 587

RESULT 7

H69378

conserved hypothetical protein AF1032 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: H69378

R.; Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Pyrococcus furiosus*

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69378

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-886 <KLE>

A;Cross-references: UNIPROT:O29230; GB:AE001032; GB:AE000782; NID:g2689355; PIDN:AAB9021

C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.2%; Score 108; DB 2; Length 886;

Best Local Similarity 28.9%; Pred. No. 3.7;

Matches 39; Conservative 22; Mismatches 28; Indels 46; Gaps 6;

QY 1 LKSIDSESDYAKGFRAPLHSLKDAKQAKLSKLELSKIDELDAEIAKLEQDLKXAVEE 49

DB 303 LRDVKEKREG-DLTREA--AGIQALQKAERDQSKLEIEITKRIEELERLEFEKSHRLLE 359

QY 50 -----AKLEQD-----LKAVERNVEDYSTGLEKTTAAKKT 82

DB 360 TLKPRMDRMQGIKAKLEKNTLTPDKVRMYDLSKAKEEKEI-----TEKLLKLI-AKKS 414

QY 83 ELEKTEADLKXAVEE 97

DB 415 SLKTRGQAKLXAVEE 429

RESULT 8

S48396

trpomyosin TPM2 - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein YJL138c

C;Species: *Saccharomyces cerevisiae*

C;Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 09-Jul-2004

C;Accession: S48396; A56490

R.; Churcher, C.

submitted to the EMBL Data Library, September 1994

A;Reference number: S48310

A;Accession: S48396

A;Molecule type: DNA

A;Residues: 1-161 <CHU>

A;Cross-references: UNIPROT:P40414; GB:Z47047; EMBL:Z38059; NID:g603997; PID:g763208; MUI

J.; Dress, B.; Brown, C.; Barrell, B.G.; Bretscher, A.

J. Cell Biol. 128, 383-392, 1995

A;Title: Tropomyosin is essential in yeast, yet the TPM1 and TPM2 products perform distir

A;Reference number: A56490; MUID:95146545; PMID:7844152

A;Accession: A56490

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-161 <DRE>

A;Cross-references: GB:Z47047; GB:Z38059; NID:g603997; PID:g763208

C;Genetics:

A;Gene: SGD:TPM2

A;Cross-references: SGD:S0001400; MIPS:YJL138c

A;Map position: 9L

C;Superfamily: tropomyosin TPM1

C;Keywords: cytoskeleton

Query Match 22.1%; Score 107.5; DB 2; Length 161;

Best Local Similarity 32.3%; Pred. No. 0.72;

Matches 31; Conservative 21; Mismatches 39; Indels 5; Gaps 2;

QY 6 ESESDYAKGFRAPLHSLKDAKQAKLSKLELSKIDELDAEIAKLEQDLKXAVEE 61

DB 14 ESESWQEKYEELREQLKEQNTKEIKLSAKNEQDSEVEKLESLQSLTKQLAED 73

QY 62 NNNVEDYSTGLEKTTAAKTELEKTEADLKXAVEE 97

DB 74 SNNLRS-NNENTYTKNQDLEQLEDSKAKLKEAMDK 108

RESULT 9

S52536

fcrA 15 protein - Streptococcus pyogenes

C;Species: Streptococcus pyogenes

C;Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004

C;Accession: S52536

R.; Katerov, V.; Schalen, C.; Totolian, A.A.

Mol. Gen. Genet. 245, 78-85, 1994

A;Title: Sequencing of genes within the vir regulon of Streptococcus pyogenes type M15 -

A;Reference number: S52535; MUID:95147851; PMID:7845360

A;Accession: S52536

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-388 <KAT>

A;Cross-references: UNIPROT:Q53474; GB:S75411; NID:g914107; PIDN:AAB33261.1; PID:g914109

C;Superfamily: M5 protein

Query Match 21.9%; Score 106.5; DB 2; Length 388;

Best Local Similarity 29.6%; Pred. No. 2;

Matches 37; Conservative 16; Mismatches 39; Indels 33; Gaps 3;

QY 1 LKSIDSESDYAKGFRAPLHSLKDAKQAKLSKLELSKIDELDAEIAKLEQDLKXAVEE 36

DB 171 LNOQDASKTEETAKLQSEAAATLENLIGSAKRELTDLQAKLDTATBAKAKLESQVTTLENL 230

QY 37 -----ELSD---KIDELDAEIAKLEQDLKXAVEENNNVEDYSTGLEKTTAAKKT 87

DB 231 LGSAKRELTDLQAKLDAANAERKAKLSQQAALQKLEATKELADLQAKLAATNQEKEL 290

QY 88 EADLK 92

DB 291 EAEAK 295

RESULT 10

A33939

Fc gamma (IgG) receptor II precursor - Streptococcus sp. (fragment)  
C;Species: Streptococcus sp.  
C;Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 26-Aug-1999  
A;Accession: A33939  
R;Heath, D.G.; Cleary, P.P.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4741-4745, 1989  
A;Title: Fc-receptor and M-protein genes of group A streptococci are products of gene du  
A;Reference number: A33939; MUID:89282846; PMID:2660147  
A;Accession: A33939  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-405 <HEA>  
A;Cross-references: GB:M22532; NID:g153628; PIDN:AA895296.1; PID:g552003  
C;Superfamily: M5 protein  
C;Keywords: immunoglobulin receptor

Query Match 21.9%; Score 106.5; DB 2; Length 405;  
Best Local Similarity 29.6%; Pred. No. 2.1;  
Matches 37; Conservative 16; Mismatches 39; Indels 33; Gaps 3;

Qy 1 LKEIDSESEDYAKEGFRA-----PLHSKLDAAQAKLSKLE----- 36  
Db 208 LKQDASKTEIAIKLQSEATLENLGSARKELTDLQAKLDATATRAEKAKLESQVTTLENL 267  
Qy 37 -----ELSD---KIDELDAEIAKLELDQKAVENNVVDYSTGLEKTIAAKKTLELEKT 87  
Db 268 LGSARKELTDLQAKLDAAAEKELQSQAAALEKLEATKKELADLQAKLAATNQEKEL 327  
Qy 88 EADLK 92  
Db 328 EAEAK 332

RESULT 11  
S70332  
outer surface protein F precursor - Lyme disease spirochete  
C;Species: Borrelia burgdorferi (lyme disease spirochete)  
C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S70532  
R;Aking, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norga  
Mol. Microbiol. 19, 507-520, 1995  
A;Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi oute  
A;Reference number: S70531; MUID:96342380; PMID:8748034  
A;Accession: S70532  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-229 <AKI>  
A;Cross-references: UNIPROT:Q44735; EMBL:U19754; NID:g3318660; PIDN:AAC26147.1; PID:g896  
C;Genetics:  
A;Gene: ospF  
C;Superfamily: outer surface protein F ospF  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-229/Product: outer surface protein F #status predicted <MAT>

Query Match 21.6%; Score 105; DB 2; Length 229;  
Best Local Similarity 30.3%; Pred. No. 1.5;  
Matches 36; Conservative 27; Mismatches 32; Indels 24; Gaps 7;

Qy 1 LKEIDSESE---SEDYAK---EGFRAPLH-----SKLDAQAKLSKLEELSDKIDELDAE 48  
Db 33 VQDLESSEQNVKTEQBIKQKQVGFLEITKDLNKLDTKEIE-KRIQELKEKIEKLEAK 91  
Qy 49 IAKL-----EDOLKAVEE---NNNVEDYSTGLEKTIAAKKTLEKTEADLKKA VNE 97  
Db 92 KTSLSKTYSEYEKLUKQIKELKKGADLED-KLKGLEDLSLKKKKEERKALEADAKKFE 149

RESULT 12  
T05409  
hypothetical protein F10M6.170 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T05409

R;Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzi, A.; N  
submitted to the Protein Sequence Database, February 1998  
A;Reference number: Z15414  
A;Accession: T05409  
A;Molecule type: DNA  
A;Residues: 1-764 <BEV>  
A;Cross-references: UNIPROT:O49371; EMBL:ALQ21811  
A;Experimental source: cultivar Columbia; BAC clone F10M6  
C;Genetics:  
A;Map position: 4  
A;Note: F10M6.170

Query Match 21.6%; Score 105; DB 2; Length 764;  
Best Local Similarity 30.7%; Pred. No. 5;  
Matches 35; Conservative 24; Mismatches 37; Indels 18; Gaps 4;

Qy 2 KEIDSESEDYAKEGFRAPLHSLKLDAAQAKLSKLELSKIDELDAEIAKLELDQKAVE 60  
Db 163 REIEELKHKLREDERAALQSSLTLEKEELEKQROEIANRSEKVSMAISEFESKQLS 222  
Qy 61 ENNV-----EDYS-----TEGLEKTIAAKKTLEK---TEADLKKA VNE 97  
Db 223 KANENVKQEGEYIALQRALEKEEELISKATKLEQEKLRTEANLKKQTEE 276

RESULT 13  
S23326  
gene ML2.2 protein precursor - Streptococcus pyogenes  
C;Species: Streptococcus pyogenes  
C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S23326  
R;Bessen, D.E.; Fischetti, V.A.  
Infect. Immun. 60, 124-135, 1992  
A;Title: Nucleotide sequences of two adjacent M or M-like protein genes of group A strep  
A;Reference number: S23325; MUID:92104662; PMID:1370269  
A;Accession: S23326  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-372 <BES>  
A;Cross-references: UNIPROT:P50469; EMBL:X61276; NID:g47369; PIDN:CAA43582.1; PID:g47371  
C;Superfamily: M5 protein

Query Match 21.5%; Score 104.5; DB 2; Length 372;  
Best Local Similarity 30.8%; Pred. No. 2.7;  
Matches 37; Conservative 17; Mismatches 37; Indels 29; Gaps 5;

Qy 3 EIDSESEDYAK-----EGFRAPLHS-----KLDAAQAKLSKLE-----ELSD 40  
Db 48 EVKESKEQYKTLALRGENADLRNVNAKYLEKINAEKKNKLEINKELNENYKLOD 107  
Qy 41 KIDELDAEIAKLELDQKAVENNVVDYSTGLEKTIAAKKT---ELE-----KTEADLKK 93  
Db 108 GIDALEKEKEDLKTTLAKTTKENEISEASRKGLSRDLASRTAKKLEAKHQKLEAKNK 167

RESULT 14  
S35760  
fcrA protein precursor - Streptococcus pyogenes  
C;Species: Streptococcus pyogenes  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S35760; A42711  
R;Podbielski, A.  
submitted to the EMBL Data Library, November 1992  
A;Reference number: S35760  
A;Accession: S35760  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-415 <POD>  
A;Cross-references: UNIPROT:Q54859; EMBL:X69324; NID:g311759; PIDN:CAA49165.1; PID:g3117  
R;Haanes, E.J.; Heath, D.G.; Cleary, P.P.  
J. Bacteriol. 174, 4967-4976, 1992  
A;Title: Architecture of the vir regulons of group A streptococci parallels opacity fact  
A;Reference number: A42711; MUID:92332431; PMID:1385809

[illegible]

Search completed: June 21, 2005, 10:11:59

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-13

Perfect score: 486

Sequence: 1 LKEIDSESEDYAKGFRAP.....KKTELEKTEADLKKA VNEPE 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | ID       | Description        |
|------------|-------|---------------|--------|----------|--------------------|
| 1          | 474   | 97.5          | 415    | 2 Q9LAY1 | Q9lay1 streptococc |
| 2          | 453   | 93.2          | 619    | 2 Q54972 | Q54972 streptococc |
| 3          | 453   | 93.2          | 619    | 2 Q8DR10 | Q8dr10 streptococc |
| 4          | 446   | 91.8          | 417    | 2 Q9LAY3 | Q9lay3 streptococc |
| 5          | 385   | 79.2          | 739    | 2 Q9ROT4 | Q9rot4 streptococc |
| 6          | 385   | 79.2          | 820    | 2 Q9ROT1 | Q9rot1 streptococc |
| 7          | 385   | 79.2          | 929    | 2 Q9KK19 | Q9kk19 streptococc |
| 8          | 385   | 79.2          | 929    | 2 Q9ZAY5 | Q9zay5 streptococc |
| 9          | 372   | 76.5          | 437    | 2 Q9LAY4 | Q9lay4 streptococc |
| 10         | 370   | 76.1          | 395    | 2 Q9LAY2 | Q9lay2 streptococc |
| 11         | 370   | 76.1          | 408    | 2 Q9LAY0 | Q9lay0 streptococc |
| 12         | 366   | 75.3          | 99     | 2 Q8KOK4 | Q8kqk4 streptococc |
| 13         | 366   | 75.3          | 249    | 2 Q9L575 | Q9l575 streptococc |
| 14         | 359   | 73.9          | 224    | 2 Q8GNS8 | Q8gns8 streptococc |
| 15         | 353   | 72.6          | 426    | 2 Q9LAY5 | Q9lay5 streptococc |
| 16         | 350.5 | 72.1          | 869    | 2 Q9KK27 | Q9kk27 streptococc |
| 17         | 339.5 | 69.9          | 246    | 2 Q9L578 | Q9l578 streptococc |
| 18         | 338.5 | 69.7          | 225    | 2 Q9L591 | Q9l591 streptococc |
| 19         | 334.5 | 68.8          | 255    | 2 Q9L581 | Q9l581 streptococc |
| 20         | 334.5 | 68.8          | 255    | 2 Q9L586 | Q9l586 streptococc |
| 21         | 332.5 | 68.4          | 222    | 2 Q9L577 | Q9l577 streptococc |
| 22         | 332.5 | 68.4          | 262    | 2 Q9L576 | Q9l576 streptococc |
| 23         | 332.5 | 68.4          | 415    | 2 Q9LAY7 | Q9lay7 streptococc |
| 24         | 329.5 | 67.8          | 416    | 2 Q9LAY8 | Q9lay8 streptococc |
| 25         | 327.5 | 67.4          | 406    | 2 Q9LAZ0 | Q9laz0 streptococc |
| 26         | 324.5 | 66.8          | 393    | 2 Q9LAZ3 | Q9laz3 streptococc |
| 27         | 321.5 | 66.2          | 394    | 2 Q9LAY6 | Q9lay6 streptococc |
| 28         | 321.5 | 66.2          | 395    | 2 Q9LAZ1 | Q9laz1 streptococc |
| 29         | 319.5 | 65.7          | 194    | 2 Q9L5B5 | Q9l5b5 streptococc |
| 30         | 319.5 | 65.7          | 218    | 2 Q6UEB2 | Q6ueb2 streptococc |
| 31         | 319.5 | 65.7          | 233    | 2 Q9L568 | Q9l568 streptococc |

|    |       |      |     |          |                    |
|----|-------|------|-----|----------|--------------------|
| 32 | 319.5 | 65.7 | 236 | 2 Q9L569 | Q9l569 streptococc |
| 33 | 319.5 | 65.7 | 243 | 2 Q9L564 | Q9l564 streptococc |
| 34 | 319.5 | 65.7 | 243 | 2 Q9L567 | Q9l567 streptococc |
| 35 | 319.5 | 65.7 | 244 | 2 Q9L565 | Q9l565 streptococc |
| 36 | 319.5 | 65.7 | 247 | 2 Q9L566 | Q9l566 streptococc |
| 37 | 319.5 | 65.7 | 249 | 2 Q9L570 | Q9l570 streptococc |
| 38 | 319.5 | 65.7 | 254 | 2 Q9L563 | Q9l563 streptococc |
| 39 | 319.5 | 65.7 | 401 | 2 Q9LAZ2 | Q9laz2 streptococc |
| 40 | 316.5 | 65.1 | 340 | 2 Q8KOK5 | Q8kqk5 streptococc |
| 41 | 313.5 | 64.5 | 207 | 2 Q8GNS9 | Q8gns9 streptococc |
| 42 | 309.5 | 63.7 | 237 | 2 Q9L592 | Q9l592 streptococc |
| 43 | 309.5 | 63.7 | 395 | 2 Q9LAY9 | Q9lay9 streptococc |
| 44 | 197.5 | 40.6 | 653 | 2 Q34097 | Q34097 streptococc |
| 45 | 179.5 | 36.9 | 246 | 2 Q9L5B4 | Q9l5b4 streptococc |

#### ALIGNMENTS

##### RESULT 1

Q9LAY1 PRELIMINARY; PRT; 415 AA.  
AC Q9LAY1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE PepA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WU2;  
RX MEDLINE=20448953; PubMed=1092499;  
DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K.; Becker R.; Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination  
in Streptococcus pneumoniae.";  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071814; AAF27710.1; -;  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON\_TER 415 415  
SQ SEQUENCE 415 AA; 46075 MW; 213C1AF7FF21642F CRC64;

Query Match 97.5%; Score 474; DB 2; Length 415;  
Best Local Similarity 98.0%; Pred. No. 7e-23; Indels 0; Gaps 0;  
Matches 97; Conservative 1; Mismatches 1;

QY 1 LKEIDSESEDYAKGFRAPLHSLDKAKLSKLELSDKIDELDAEIAKLEDLQKAVE 60  
|||||  
Db 221 LKEIDSESEDYAKGFRAPLQSKLDKAKLSKLELSDKIDELDAEIAKLEDLQKAVE 280  
|||||  
QY 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKA VNEPE 99  
|||||  
Db 281 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKA VNEPE 319  
|||||

##### RESULT 2

Q54972 PRELIMINARY; PRT; 619 AA.  
AC Q54972;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Pneumococcal surface protein A precursor.  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92105030; PubMed=1729249;
RA  Yother J., Briles D.E.;
RT  "Structural properties and evolutionary relationships of PspA, a
RT  surface protein of Streptococcus pneumoniae, as revealed by sequence
RT  analysis.";
RL  J. Bacteriol. 174:601-609(1992).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Yother J., Briles D.E.;
RL  EMBL; M74122; AAA27018.1; -.
DR  PIR; A41971; A41971.
DR  PIR; A97887; A97887.
DR  HSP; P06653; IHCX.
DR  InterPro; IPR002479; CW binding.
DR  InterPro; IPR002345; Lipocalin.
DR  Pfam; PF01473; CW_binding_1; 10.
DR  PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW  Signal.
FT  SIGNAL 1 31 Potential.
FT  CHAIN 32 619 pneumococcal surface protein A.
SQ  SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.2e-21;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKDIDELDAEIAKLEDLQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAAKAKLSKLELSKDIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321

RESULT 3
Q8DR10 ID Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name=pspA; OrderedLocNames=sp0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
RA Glass J.L.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSP; P06653; IHCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.

```

```

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 93.2%; Score 453; DB 2; Length 619;
Best Local Similarity 93.9%; Pred. No. 2.2e-21;
Matches 93; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKDIDELDAEIAKLEDLQKAAE 60
Db 223 LKEIDSESDYAKGFRAPLQSKLDAAKAKLSKLELSKDIDELDAEIAKLEDLQKAAE 282

Qy 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 283 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 321

RESULT 4
Q9LAY3 ID Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSP; P00192; 256B.
DR NON_TER 417 417
FT NON_TER 417 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506BEC CRC64;

Query Match 91.8%; Score 446; DB 2; Length 417;
Best Local Similarity 91.9%; Pred. No. 4.4e-21;
Matches 91; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSKDIDELDAEIAKLEDLQKAAE 60
Db 213 LKEIDSESDYVKEGFRAPLQSELDAAKQAKLSKLELSKDIDELDAEIAKLEDLQKAAE 272

Qy 61 ENNVEDYSTEGLEKTIAAKTELEKTEADLKKAVNEPE 99
Db 273 ENNVEDYFKEGLEKTIAAKAELEKTEADLKKAVNEPE 311

RESULT 5
Q9RQT4 ID Q9RQT4 PRELIMINARY; PRT; 739 AA.
AC Q9RQT4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=E134;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL: AF068647; AAF13457.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR005877; Gpos YSIRK.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW binding_1; 1.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; YSIRK signal; 1.
DR TIGRFAMs: TIGR01168; YSIRK signal; 1.
KW Hypothetical protein.
FT NON TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EE2F2F676ABF989 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 739;
Best Local Similarity 81.8%; Pred. No. 5.9e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDOLKAVE 60
DB 537 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDOLKAVE 596
QY 61 ENNVNVEDYSTEGLEKTTAAKTELEKTEADLKAVNEPE 99
DB 597 GNNVVEAYFKEGLEKTTAAKTELEKTEADLKAVNEPE 635

RESULT 6
Q9RQT1 ID Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Names: pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9183;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL: AF068650; AAF13460.1; -.
DR HSSP: P04268; IIC2.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR005877; Gpos YSIRK.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW binding_1; 1.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; YSIRK signal; 1.
DR TIGRFAMs: TIGR01168; YSIRK signal; 1.
KW Hypothetical protein.
FT NON TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 820;
Best Local Similarity 81.8%; Pred. No. 6.5e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 1 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDOLKAVE 60
DB 530 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDOLKAVE 589
QY 61 ENNVNVEDYSTEGLEKTTAAKTELEKTEADLKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTTAAKTELEKTEADLKAVNEPE 628

RESULT 7
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Names: pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=srf10;
RX MEDLINE=21886621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL: AF154037; AAF73809.1; -.
DR HSSP: P06653; IH8G.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR005877; Gpos YSIRK.
DR InterPro: IPR009053; Prefoldin.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW binding_1; 11.
DR Pfam: PF05062; RICH; 2.
DR Pfam: PF04650; YSIRK signal; 1.
DR TIGRFAMs: TIGR01168; YSIRK signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 929;
Best Local Similarity 81.8%; Pred. No. 7.3e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDOLKAVE 60
DB 530 LXEIDSESDYAKGFRAPLHSLDKDAKQAKLSKLELSDKIDELDAETAKLEDOLKAVE 589
QY 61 ENNVNVEDYSTEGLEKTTAAKTELEKTEADLKAVNEPE 99
DB 590 GNNVVEAYFKEGLEKTTAAKTELEKTEADLKAVNEPE 628

RESULT 8
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Names: pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;

```

```

RX MEDLINE=20038319; PubMed=10569772;
RA Brooke-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
protein, PspC, which elicits cross-reactive antibodies to PspA and
provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 79.2%; Score 385; DB 2; Length 929;
Best Local Similarity 81.8%; Pred. No. 7.3e-17;
Matches 81; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 530 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 589
Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 590 GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 9
Q9LAY4
ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYSIN.
FT NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 76.5%; Score 372; DB 2; Length 437;
Best Local Similarity 79.8%; Pred. No. 2.5e-16;
Matches 79; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 235 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 294
Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 295 GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 333

us-10-674-755-13.rup

```

```

RESULT 10
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E56796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYSIN.
FT NON TER 395
SQ SEQUENCE 395 AA; 42963 MW; 58B6EF956BCBCC1E CRC64;

Query Match 76.1%; Score 370; DB 2; Length 395;
Best Local Similarity 78.8%; Pred. No. 3.1e-16;
Matches 78; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 1 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 60
Db 225 LKEIDSESDYAKGFRAPLHSLDKAKQAKLSKLELSKIDELDAEIAKLELDQKAVE 284
Qy 61 ENNVEDYSTEGLEKTTAAKTELEKTEADLKKAVNEPE 99
Db 285 GNNVVEAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 323

us-10-674-755-13.rup

```



```

RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 1
FT NON_TER 249
FT NON_TER 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 75.3%; Score 366; DB 2; Length 249;
Best Local Similarity 77.8%; Pred. No. 3.7e-16;
Matches 77; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Qy 1 LKIDSESDYAKGFRAPLHSLDKAKAKLSKLEELSDKIDELDAEIAKLELDQKAVE 60
Db 74 LKIDSDSDYIKGFRAPLQSELDTKAKULKLELSKGIIEELDAEIAEVLQKDAE 133

Qy 61 ENNVEDYSPGELKTIAAKTELEKTEADLKAVNEPE 99
Db 134 GNNVYAFKGELEKTTAEKKAELKAEADLKAVDEPE 172

RESULT 14
Q8GNS8 PRELIMINARY; PRT; 224 AA.
ID Q8GNS8
AC Q8GNS8;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicuonzo G., Gherardi S., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia G., Pantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
RT Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AA37735.1; -.
DR HSP; P00192; IAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 224
FT NON_TER 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 73.9%; Score 359; DB 2; Length 224;
Best Local Similarity 75.8%; Pred. No. 9.3e-16;
Matches 75; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

Qy 1 LKIDSESDYAKGFRAPLHSLDKAKAKLSKLEELSDKIDELDAEIAKLELDQKAVE 60
Db 17 LKIDNEDSDYKGFAPLQSELDTKAKULKLELSKGIIEELDAEIAEVLQKDAE 76

Qy 61 ENNVEDYSPGELKTIAAKTELEKTEADLKAVNEPE 99
Db 77 GNNVYAFKGELEKTTAEKKAELKAEADLKAVDEPE 115

RESULT 15
Q9LAY5 PRELIMINARY; PRT; 426 AA.
ID Q9LAY5
AC Q9LAY5;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBL5;  
RX MEDLINE=20448953; PubMed=10992499;  
RA DOI=10.1128/IAI.68.10.5889-5900.2000; D.E.;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of PspA: mosaic genes and evidence for past recombination  
in Streptococcus pneumoniae";  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071810; AAP27706.1; -.  
DR HSP; P00192; IMGT.  
DR InterPro; IPR011047; Quin\_alc\_DH\_like.  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON\_TER 426  
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;  
  
Query Match 72.6%; Score 353; DB 2; Length 426;  
Best Local Similarity 74.7%; Pred. No. 4e-15;  
Matches 74; Conservative 8; Mismatches 17; Indels 0; Gaps 0;  
  
Qy 1 LKEIDSESDYAKGFRAPLHSLDKAQAKLSKLELSDKIDELDAEIAKLEDLKAVE 60  
Db 215 LKQINESDSEDYVKEGLRAPLQSELDTKAKLLKLELSGKIEELDAEIAELEVQLKDAE 274  
  
Qy 61 ENNVVDYSTGLEKTTAAKTELEKTEADLKKA VNEPE 99  
Db 275 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKA VDEPE 313

Search completed: June 21, 2005, 10:22:11  
Job time : 62.3194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKDIIDSESDYAKGERAP.....KKAELEKARADLKKAVIDPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID      | Description        |
|------------|-------|-------------|--------|------------|--------------------|
| 1          | 472   | 97.9        | 206    | 2 AAW14574 | Aaw14574 Streptoco |
| 2          | 472   | 97.9        | 206    | 7 ABW02608 | Abw02608 Db15c pne |
| 3          | 472   | 97.9        | 8991   | 6 ABU08487 | Abu08487 S. pneumo |
| 4          | 440   | 91.3        | 170    | 7 ABW02614 | Abw02614 Rct135c p |
| 5          | 440   | 91.3        | 181    | 7 ABW02596 | Abw02596 0922134c  |
| 6          | 440   | 91.3        | 865    | 6 ABU08489 | Abu08489 S. pneumo |
| 7          | 440   | 91.3        | 929    | 2 AAW14593 | Aaw14593 Streptoco |
| 8          | 440   | 91.3        | 929    | 2 AAY43384 | Aay43384 S. pneumo |
| 9          | 437   | 90.7        | 188    | 2 AAW14580 | Aaw14580 Streptoco |
| 10         | 437   | 90.7        | 188    | 7 ABW02613 | Abw02613 Rct129c p |
| 11         | 428   | 88.8        | 588    | 6 ABU08491 | Abu08491 Coiled co |
| 12         | 428   | 88.8        | 589    | 2 AAY43392 | Aay43392 PspC alph |
| 13         | 426   | 88.4        | 204    | 2 AAW14578 | Aaw14578 Streptoco |
| 14         | 426   | 88.4        | 204    | 7 ABW02612 | Abw02612 Rct123c p |
| 15         | 425.5 | 88.3        | 180    | 2 AAW14562 | Aaw14562 Streptoco |
| 16         | 422.5 | 87.7        | 187    | 2 AAW14579 | Aaw14579 Streptoco |
| 17         | 405   | 84.0        | 1231   | 6 ABU08490 | Abu08490 Fragment  |
| 18         | 389   | 80.7        | 198    | 2 AAW14581 | Aaw14581 Streptoco |
| 19         | 386   | 80.1        | 198    | 7 ABW02615 | Abw02615 Rx1c pneu |
| 20         | 386   | 80.1        | 204    | 2 AAW14571 | Aaw14571 Streptoco |
| 21         | 386   | 80.1        | 204    | 7 ABW02605 | Abw02605 Bf1019c p |
| 22         | 386   | 80.1        | 315    | 2 AAY04375 | Aay04375 Streptoco |
| 23         | 386   | 80.1        | 619    | 2 AAR63437 | Aar63437 Pneumococ |
| 24         | 386   | 80.1        | 619    | 2 AAR87598 | Aar87598 Pneumococ |
| 25         | 386   | 80.1        | 619    | 2 AAR86911 | Aar86911 Pneumococ |

RESULT 1

AAW14574

ID AAW14574 standard; protein; 206 AA.

XX

AC AAW14574;

XX

DT 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX

DE Streptococcus pneumoniae PspA central region.

XX

KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;

KW bacteraemia; pneumonia.

XX

OS Streptococcus pneumoniae; strain Db15.

XX

FH Key Location/Qualifiers

FT Misc-difference 50

FT /note= "unidentified amino acid"

XX

PN WO9709994-A1.

XX

PD 20-MAR-1997.

XX

PF 16-SEP-1996; 96WO-US014819.

XX

PR 15-SEP-1995; 95US-00529055.

XX

PA (UABR-) UAB RES FOUND.

XX

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX

DR WPI; 1997-202002/18.

XX

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used

PT in vaccines for protecting animals against S.pneumoniae infection.

XX

PS Example 6; Fig 13; 296pp; English.

XX

CC This sequence shows the central portion, including the C-terminus of the

CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Db15.

CC Comparison of the N-terminal and central regions (AAW14533-57 and

CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can

CC be used to divide the strains into several families based on sequence

CC homologies. PspA polypeptides, or fragments of them, can be used in

CC vaccines to protect animals against S. pneumoniae infection and hence for

CC



Query Match 97.9%; Score 472; DB 6; Length 8991;  
Best Local Similarity 99.0%; Pred. No. 2.2e-35;  
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60  
||:|||||  
DB 5888 LKIDSDSDYAKGERAPLOSELDTKAKLLKLEELSGKIEELDAEIXEVLQKDAE 5947  
||:|||||

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
||:|||||  
DB 5948 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 5986  
||:|||||

RESULT 4  
ABW02614  
ID ABW02614 standard; protein; 170 AA.  
XX  
AC ABW02614;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Rct135c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
PN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX WPI; 2003-862841/80.  
XX

Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

XX Example 6; SEQ ID NO 60; 121pp; English.

XX The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspAs) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as vaccines and in gene therapy. The present sequence is Rct135c pneumococcal surface protein A (PspA) central region. This sequence is used in the exemplification of the invention

XX Sequence 170 AA;

Query Match 91.3%; Score 440; DB 7; Length 170;  
Best Local Similarity 90.9%; Pred. No. 2.1e-34;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60  
||:|||||  
DB 5888 LKIDSDSDYAKGERAPLOSELDTKAKLLKLEELSGKIEELDAEIXEVLQKDAE 5947  
||:|||||

Db 1 LKEIDSDSDYKGLRAPLOSKLDTKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60  
QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
||:|||||  
Db 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
||:|||||

RESULT 5  
ABW02596  
ID ABW02596 standard; protein; 181 AA.  
XX  
AC ABW02596;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE 0922134c pneumococcal surface protein A (PspA) central region.  
XX  
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KW immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
PN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.  
XX  
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX WPI; 2003-862841/80.  
XX

Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

XX Example 6; SEQ ID NO 42; 121pp; English.

XX The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspAs) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as vaccines and in gene therapy. The present sequence is 0922134c pneumococcal surface protein A (PspA) central region. This sequence is used in the exemplification of the invention

XX Sequence 181 AA;

Query Match 91.3%; Score 440; DB 7; Length 181;  
Best Local Similarity 90.9%; Pred. No. 2.2e-34;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60  
||:|||||  
Db 1 LKEIDSDSDYKGLRAPLOSKLDTKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60  
||:|||||

QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
||:|||||  
Db 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
||:|||||

```

RESULT 6
ABU08489
ID ABU08489 standard; protein; 865 AA.
XX AC ABU08489;
XX DT 24-JUN-2003 (first entry)
XX DE S. pneumoniae pneumococcal surface protein C (PspC) protein.
XX KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
XX KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
XX KW antibacterial.
XX OS Streptococcus pneumoniae.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Signal_peptide
XX FT /label= Signal_peptide
XX FT /label= Mature_PspC_protein
XX US6500613-B1.
XX 31-DEC-2002. 96US-00714741.
XX 16-SEP-1996; 96US-00714741.
XX 15-SEP-1995; 95US-00529055.
XX (UVAL-) UNIV ALABAMA.
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 2003-361534/34.
XX N-PSDB; ABX95377.
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
XX hybridization probe, comprises pneumococcal surface protein having alpha-
XX helical, proline rich and repeat regions.
XX Claim 3; Fig 21; 186pp; English.
XX The present invention relates to the isolation of Streptococcus
XX pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX like protein having alpha-helical, proline rich and repeat regions. The
XX PspC and PspA proteins may be used in a vaccine to protect against
XX pneumococcal infections. The polynucleotide sequences encoding PspC and
XX PspA may be used for the expression of the proteins, and as PCR primers
XX or hybridisation probes. The present sequence represents S. pneumoniae
XX PspC protein
XX SQ Sequence 865 AA;
Query Match 91.3%; Score 440; DB 6; Length 865;
Best Local Similarity 90.9%; Pred. No. 1.5e-33;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 1 LKIDDESDSDYAKGERAPLQSLDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60
Db 466 LKEIDSDSDYKLEGLRAPLQSLDTKKAKLLKLEELSDKIDELDAEIXEVLQKDAE 525
Qy 61 GNNVVEYFKGLEKTTAEKKAELKAPADLKKAVIDEPE 99
Db 526 GNNVVEYFKGLEKTTAEKKAELKAPADLKKAVIDEPE 564
RESULT 7
AAW14593
ID AAW14593 standard; protein; 929 AA.

```

```

XX AAW14593;
XX AC 17-OCT-2003 (revised)
XX DT 27-OCT-1997 (first entry)
XX DE Streptococcus pneumoniae PspC.
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX KW bacteraemia; pneumonia.
XX OS Streptococcus pneumoniae; strain EF6796.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Sig_peptide
XX FT /label= Mat_protein
XX FT /label= Alpha-helix
XX FT /label= Repeat_1
XX FT /note= "alpha-helical repeat region"
XX FT /label= Coiled-coil
XX FT /note= "breaks in 7-residue periodicity of the coiled-
XX FT coil occur at amino acids 136-141, 261-304 and 383-387"
XX FT /label= Repeat_2
XX FT /note= "alpha-helical repeat region"
XX FT /label= Proline-rich
XX FT /label= C-terminal
XX WO9709994-A1.
XX 20-MAR-1997.
XX 16-SEP-1996; 96WO-US014819.
XX 15-SEP-1995; 95US-00529055.
XX (UABR-) UAB RES FOUND.
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 1997-202002/18.
XX N-PSDB; AAT61728.
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX Disclosure; Fig 13; 296pp; English.
XX This sequence comprises the pneumococcal protein surface C (pspc) of
XX Streptococcus pneumoniae strain EF6796. The sequence was deduced from the
XX pspC gene (AAT61728). Like PspA, PspC has an alpha-helical coiled-coil
XX region, proline-rich central region, repeat regions, with a choline
XX binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides
XX share 3 regions of high sequence identity. One is a protection-eliciting
XX region present within the alpha-helical domain. The others are the
XX proline-rich domain and a repeat domain shared with other choline-binding
XX proteins and thought to play a role in cell surface association. PspC and
XX PspA polypeptides, and their fragments, can be used in vaccines to
XX protect against S. pneumoniae infection and hence for the prevention of
XX diseases such as otitis media, meningitis, bacteraemia and pneumonia.
XX (Updated on 17-OCT-2003 to standardise OS field)
XX SQ Sequence 929 AA;
Query Match 91.3%; Score 440; DB 2; Length 929;

```

Best Local Similarity 90.9%; Pred. No. 1.6e-33; Mismatches 4; Indels 0; Gaps 0;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIBELDAEIXELEVQLKDAE 60  
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSGKIBELDAEIXELEVQLKDAE 589

QY 61 GNNVEAYFKGLEKTTAEKAELEKAEADLKKAVDEPE 99  
DB 590 GNNVEAYFKGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 8  
ID AAY43384 standard; protein; 929 AA.  
XX AAY43384;  
XX 27-JAN-2000 (first entry)  
XX S. pneumoniae PspC protein sequence.  
XX PspC gene; pneumococcal surface protein C; epitope; diagnosis;  
KW epitopic region; immunogenic composition; vaccine composition; therapy;  
KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;  
KW anti-PspA antibody.  
XX Streptococcus pneumoniae.  
OS WO9953940-A1.  
PN 28-OCT-1999.  
XX 23-APR-1999; 99WO-US008895.  
XX 23-APR-1998; 98US-0082728P.  
XX (UYAL-) UNIV ALABAMA.  
XX Briles DE, Hollingshead SK, Brooks-Walter A;  
PI WPI; 1999-620581/53.  
DR N-PSDB; AAZ31956.  
XX New epitope polypeptides of Pneumococcal surface protein C, used to  
PT develop products for immunological, immunogenic or vaccine compositions,  
PT particularly against Streptococcus pneumoniae infections.  
XX Example; Fig 11; 109pp; English.  
XX This sequence is the Streptococcus pneumoniae pneumococcal surface  
CC protein C. The invention relates to an isolated and/or purified  
CC polypeptide comprising at least one epitope or epitopic region of  
CC pneumococcal surface protein C (PspC). The polypeptides or vectors  
CC containing sequence encoding them can be used as immunogenic,  
CC immunological or vaccine compositions. The compositions can be used for  
CC eliciting an immunological response against Streptococcus pneumoniae  
CC (SP), which can cause otitis media, meningitis, bacteraemia and  
CC pneumonia. They can be used for eliciting an anti-PspA antibody. The  
CC nucleic acid molecules can also be used for detecting pspC, pspA or SP  
XX Sequence 929 AA;  
SQ Query Match 91.3%; Score 440; DB 2; Length 929;  
Best Local Similarity 90.9%; Pred. No. 1.6e-33;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIBELDAEIXELEVQLKDAE 60  
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSGKIBELDAEIXELEVQLKDAE 589

QY 61 GNNVEAYFKGLEKTTAEKAELEKAEADLKKAVDEPE 99  
DB 590 GNNVEAYFKGLEKTTAEKAELEKAEADLKKAVDEPE 628

Db 590 GNNVEAYFKGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 9  
AAW14580  
ID AAW14580 standard; protein; 188 AA.  
XX AAW14580;  
XX 17-OCT-2003 (revised)  
DT 28-OCT-1997 (first entry)  
XX Streptococcus pneumoniae PspA central region.  
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
KW bacteraemia; pneumonia.  
XX Streptococcus pneumoniae; strain Rct135.  
OS WO9709994-A1.  
PN 20-MAR-1997.  
XX 16-SEP-1996; 96WO-US014819.  
XX 15-SEP-1995; 95US-00529055.  
XX (UABR-) UAB RES FOUND.  
XX Briles DE, Medaniel LS, Swiatlo E, Yother J, Crain MJ;  
PI Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI; 1997-202002/18.  
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX Example 6; Fig 13; 296pp; English.  
XX This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX Sequence 188 AA;  
SQ Query Match 90.7%; Score 437; DB 2; Length 188;  
Best Local Similarity 89.9%; Pred. No. 4.5e-34;  
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLKLELSGKIBELDAEIXELEVQLKDAE 60  
DB 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSGKIBELDAEIXELEVQLKDAE 60

QY 61 GNNVEAYFKGLEKTTAEKAELEKAEADLKKAVDEPE 99  
DB 590 GNNVEAYFKGLEKTTAEKAELEKAEADLKKAVDEPE 99

RESULT 10  
ABW02613  
ID ABW02613 standard; protein; 188 AA.  
XX ABW02613;  
AC ABW02613;

XX 12-FEB-2004 (first entry)  
 XX Rct129c pneumococcal surface protein A (PspA) central region.  
 DE  
 XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 KW  
 XX Unidentified.  
 OS  
 XX US6592876-B1.  
 PN  
 XX 15-JUL-2003.  
 PD  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 XX 20-APR-1993; 93US-00048896.  
 PR  
 XX 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI  
 XX MPI; 2003-862841/80.  
 DR  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain.  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 XX Example 6; SEQ ID NO 59; 121pp; English.  
 PS  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Rct129c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 SQ Sequence 188 AA;  
 Query Match 90.7%; Score 437; DB 7; Length 188;  
 Best Local Similarity 89.9%; Pred. No. 4.5e-34;  
 Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 LKIDESDSYAKGERAPLOSELDTTKAKLKLSELSGKIBELDAEIXEVLQKDAE 60  
 Db 1 LKIDESDSYDLKEGLRAPLQSKLDTKAKLKLSELSGKIBELDAEIXEVLQKDAE 60  
 Qy 61 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKADEPE 99  
 Db 61 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKADEPD 99  
 RESULT 11  
 ABU08491  
 ID ABU08491 standard; protein; 588 AA.  
 AC  
 XX ABU08491;  
 XX  
 XX 24-JUN-2003 (first entry)  
 DT  
 XX Coiled coil motif of alpha-helix of S. pneumoniae PspC protein.  
 DE  
 XX Pneumococcal surface protein C; PspC, pneumococcal surface protein A;  
 KW

KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 XX US6500613-B1.  
 PN  
 XX 31-DEC-2002.  
 PD  
 XX 16-SEP-1996; 96US-00714741.  
 PP  
 XX 15-SEP-1995; 95US-00529055.  
 PR  
 XX (UYAL-) UNIV ALABAMA.  
 PA  
 XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI  
 XX Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 XX MPI; 2003-361534/34.  
 DR  
 XX Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.  
 PT  
 XX Disclosure; Fig 23; 186pp; English.  
 PS  
 XX The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents a coiled coil  
 CC motif of the alpha-helix of S. pneumoniae PspC protein  
 XX  
 SQ Sequence 588 AA;  
 Query Match 88.8%; Score 428; DB 6; Length 588;  
 Best Local Similarity 90.7%; Pred. No. 1.3e-32;  
 Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 LKIDESDSYAKGERAPLOSELDTTKAKLKLSELSGKIBELDAEIXEVLQKDAE 60  
 Db 492 LKIDESDSYDLKEGLRAPLQSKLDTKAKLKLSELSGKIBELDAEIXEVLQKDAE 551  
 Qy 61 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKADE 97  
 Db 552 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKADE 588  
 RESULT 12  
 AAY43392  
 ID AAY43392 standard; protein; 589 AA.  
 XX  
 AC AAY43392;  
 XX  
 XX 27-JAN-2000 (first entry)  
 DT  
 XX PspC alpha-helix coiled-coil region.  
 DE  
 XX  
 XX PspC gene; pneumococcal surface protein C; epitope; diagnosis;  
 KW epitopic region; immunogenic composition; vaccine composition; therapy;  
 KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;  
 KW anti-PspA antibody.  
 XX  
 OS Streptococcus pneumoniae.  
 OS  
 XX WO9953940-A1.  
 PN  
 XX 28-OCT-1999.  
 PD  
 XX 23-APR-1999; 99WO-US008895.  
 PF



XX PR 23-APR-1998; 98US-0082728P.  
 XX PA (UYAL-) UNIV ALABAMA.  
 XX PI Briles DE, Hollingshead SK, Brooks-Walter A;  
 XX PT WPI; 1999-620581/53.  
 XX DR  
 XX PT New epitope polypeptides of Pneumococcal surface protein C, used to  
 PT develop products for immunological, immunogenic or vaccine compositions,  
 PT particularly against Streptococcus pneumoniae infections.  
 XX PS Example 1; Fig 3; 109pp; English.  
 XX CC This sequence is the coiled-coil region of the Streptococcus pneumoniae  
 CC pneumococcal surface protein C. The invention relates to an isolated  
 CC and/or purified polypeptide comprising at least one epitope or epitopic  
 CC region of Pneumococcal surface protein C (PspC). The polypeptides or  
 CC vectors containing sequence encoding them can be used as immunogenic,  
 CC immunological or vaccine compositions. The compositions can be used for  
 CC eliciting an immunological response against Streptococcus pneumoniae  
 CC (SP), which can cause otitis media, meningitis, bacteraemia and  
 CC pneumonia. They can be used for eliciting an anti-PspA antibody. The  
 CC nucleic acid molecules can also be used for detecting pspC, pspA or SP  
 XX SQ Sequence 589 AA;  
 Query Match 88.8%; Score 428; DB 2; Length 589;  
 Best Local Similarity 90.7%; Pred. No. 1.3e-32;  
 Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLEELSGKIELDAEIXEVLKDAE 60  
 Db 493 LKIDSDSDYAKGERAPLOSELDTKKAKLKLEELSGKIELDAEIXEVLKDAE 552  
 QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKAVDE 97  
 Db 553 GNNVYAYFKEGLEKTTAEKKAELKAEADLKAVDE 589  
 RESULT 13  
 AAW14578  
 ID AAW14578 standard; protein; 204 AA.  
 XX AC AAW14578;  
 XX DT 17-OCT-2003 (revised)  
 DE 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX Streptococcus pneumoniae; strain Rct123.  
 XX Key Location/Qualifiers  
 FH Misc-difference 4 /note= "unidentified amino acid"  
 FT Misc-difference 8 /note= "unidentified amino acid"  
 FT WO9705994-A1.  
 XX 20-MAR-1997.  
 XX 16-SEP-1996; 96WO-US014819.  
 XX 15-SEP-1995; 95US-00529055.  
 XX (UABR-) UAB RES FOUND.  
 XX PA Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.

PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 XX Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX PS Example 6; Fig 13; 296pp; English.  
 XX CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX SQ Sequence 204 AA;  
 Query Match 88.4%; Score 426; DB 2; Length 204;  
 Best Local Similarity 87.9%; Pred. No. 5.7e-33;  
 Matches 87; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 LKIDSDSDYAKGERAPLOSELDTKKAKLKLEELSGKIELDAEIXEVLKDAE 60  
 Db 1 IXEXDSXSEDYKLEGLRAPLOSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60  
 QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKAVDE 99  
 Db 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKAVDE 99  
 RESULT 14  
 ABW02612  
 ID ABW02612 standard; protein; 204 AA.  
 XX AC ABW02612;  
 XX DT 12-FEB-2004 (first entry)  
 DE Rct123c pneumococcal surface protein A (PspA) central region.  
 XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX Unidentified.  
 XX Key Location/Qualifiers  
 FH Misc-difference 1.204 /note= "Xaa = Unknown amino acid"  
 FT US6592876-B1.  
 XX 15-JUL-2003.  
 XX 15-SEP-1995; 95US-00529055.  
 XX 20-APR-1993; 93US-00048896.  
 XX 06-JUN-1995; 95US-00465746.  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKIDISDSYAKGERAP.....KKALEKAEADLKKAVDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 480   | 99.6        | 99     | 4     | US-09-147-875A-14 |
| 2          | 472   | 97.9        | 206    | 4     | US-08-529-055-54  |
| 3          | 472   | 97.9        | 8931   | 4     | US-08-714-741-32  |
| 4          | 458   | 95.0        | 99     | 2     | US-08-710-749-13  |
| 5          | 451   | 93.6        | 99     | 4     | US-09-147-875A-15 |
| 6          | 448   | 92.9        | 99     | 4     | US-09-147-875A-16 |
| 7          | 440   | 91.3        | 170    | 4     | US-08-529-055-60  |
| 8          | 440   | 91.3        | 181    | 4     | US-08-529-055-42  |
| 9          | 440   | 91.3        | 864    | 4     | US-08-714-741-40  |
| 10         | 437   | 90.7        | 99     | 2     | US-08-710-749-14  |
| 11         | 437   | 90.7        | 188    | 4     | US-08-529-055-59  |
| 12         | 430   | 89.2        | 99     | 2     | US-08-710-749-17  |
| 13         | 429.5 | 89.1        | 100    | 4     | US-09-147-875A-10 |
| 14         | 428   | 88.8        | 141    | 4     | US-09-286-981B-2  |
| 15         | 428   | 88.8        | 588    | 4     | US-08-714-741-42  |
| 16         | 426   | 88.4        | 99     | 2     | US-08-710-749-15  |
| 17         | 426   | 88.4        | 204    | 4     | US-08-529-055-58  |
| 18         | 405   | 84.0        | 1231   | 4     | US-08-714-741-41  |
| 19         | 386   | 80.1        | 99     | 2     | US-08-710-749-10  |
| 20         | 386   | 80.1        | 99     | 2     | US-08-710-749-11  |
| 21         | 386   | 80.1        | 99     | 4     | US-09-147-875A-11 |
| 22         | 386   | 80.1        | 198    | 4     | US-08-529-055-61  |
| 23         | 386   | 80.1        | 204    | 4     | US-08-529-055-51  |
| 24         | 386   | 80.1        | 619    | 1     | US-08-465-746-2   |
| 25         | 386   | 80.1        | 619    | 1     | US-08-214-164-2   |
| 26         | 386   | 80.1        | 619    | 2     | US-08-467-852A-3  |
| 27         | 386   | 80.1        | 619    | 2     | US-08-246-636-2   |

|    |       |      |     |   |                   |                    |
|----|-------|------|-----|---|-------------------|--------------------|
| 28 | 386   | 80.1 | 619 | 2 | US-08-247-491A-3  | Sequence 3, Appli  |
| 29 | 386   | 80.1 | 619 | 2 | US-08-319-795-2   | Sequence 2, Appli  |
| 30 | 386   | 80.1 | 619 | 2 | US-08-468-985-2   | Sequence 2, Appli  |
| 31 | 386   | 80.1 | 619 | 3 | US-08-312-949-2   | Sequence 2, Appli  |
| 32 | 386   | 80.1 | 648 | 1 | US-08-072-070-2   | Sequence 2, Appli  |
| 33 | 386   | 80.1 | 648 | 1 | US-08-469-434-2   | Sequence 2, Appli  |
| 34 | 386   | 80.1 | 648 | 1 | US-08-214-222-2   | Sequence 2, Appli  |
| 35 | 386   | 80.1 | 648 | 2 | US-08-467-852A-2  | Sequence 2, Appli  |
| 36 | 386   | 80.1 | 648 | 2 | US-08-468-718-2   | Sequence 2, Appli  |
| 37 | 386   | 80.1 | 648 | 2 | US-08-247-491A-2  | Sequence 2, Appli  |
| 38 | 386   | 80.1 | 648 | 3 | US-08-446-201-3   | Sequence 23, Appli |
| 39 | 386   | 80.1 | 695 | 1 | US-08-127-499A-23 | Sequence 23, Appli |
| 40 | 386   | 80.1 | 695 | 1 | US-08-482-847-23  | Sequence 23, Appli |
| 41 | 374   | 77.6 | 288 | 3 | US-08-312-949-4   | Sequence 4, Appli  |
| 42 | 374   | 77.6 | 288 | 3 | US-08-446-201-4   | Sequence 4, Appli  |
| 43 | 372.5 | 77.3 | 100 | 4 | US-09-147-875A-12 | Sequence 12, Appli |
| 44 | 365   | 75.7 | 193 | 4 | US-08-529-055-49  | Sequence 49, Appli |
| 45 | 363.5 | 75.4 | 289 | 1 | US-08-072-070-4   | Sequence 4, Appli  |

#### ALIGNMENTS

RESULT 1

US-09-147-875A-14

; Sequence 14, Application US/09147875A

; Patent No. 6638516

; GENERAL INFORMATION:

; APPLICANT: BECKER et al.

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/09/147,875A

; CURRENT FILING DATE: 1999-05-24

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (1)..(99)

; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid

US-09-147-875A-14

Query Match 99.6%; Score 480; DB 4; Length 99;

Best Local Similarity 100.0%; Pred. No. 5.4e-40;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKIDISDSYAKGERAPLQSELDTKKAKLKLKLELKGKIELDABYELVQLKDAE 60

Db 1 LKIDISDSYAKGERAPLQSELDTKKAKLKLKLELKGKIELDABYELVQLKDAE 60

QY 61 GNNVAYFKEGLEKTTAKKAELEKAEADLKKAVDEPE 99

Db 61 GNNVAYFKEGLEKTTAKKAELEKAEADLKKAVDEPE 99

RESULT 2

US-08-529-055-54

; Sequence 54, Application US/08529055

; Patent No. 6592876

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; TITLE OF INVENTION: Pneumococcal Genes, Portions

; TITLE OF INVENTION: Thereof, Expression Products

; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,

; TITLE OF INVENTION: Portions and Products

; NUMBER OF SEQUENCES: 73

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-54

Query Match          97.9%; Score 472; DB 4; Length 206;
Best Local Similarity 99.0%; Pred. No. 7.8e-39;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDIDESDYAKEGRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60
Db 1 LKIDIDESDYAKEGLRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60

Qy 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 3
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 650613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match          97.9%; Score 472; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 6.6e-37;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDIDESDYAKEGRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 60
Db 5888 LKIDIDESDYAKEGLRAPLQSELDTKKAKLLKLELSGKIEELDAAIXELEVQLKDAE 5947

Qy 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 5948 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 5986

RESULT 4
US-08-710-749-13
; Sequence 13, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
```

TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-13

Query Match 95.0%; Score 458; DB 2; Length 99;  
Best Local Similarity 96.0%; Pred. No. 7.7e-38;  
Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60  
DB 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60  
QY 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99  
DB 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99

RESULT 5  
US-09-147-875A-15  
; Sequence 15, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-15

Query Match 93.6%; Score 451; DB 4; Length 99;  
Best Local Similarity 92.9%; Pred. No. 3.7e-37;  
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60  
DB 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60  
QY 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99  
DB 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99

RESULT 6  
US-09-147-875A-16  
; Sequence 16, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-16

Query Match 92.9%; Score 448; DB 4; Length 99;  
Best Local Similarity 91.9%; Pred. No. 7.3e-37;  
Matches 91; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60  
DB 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60

QY 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99  
DB 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99  
RESULT 7  
US-08-529-055-60  
; Sequence 60, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; THEREOF, Expression Products  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Thereof, Expression Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 170 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-60

Query Match 91.3%; Score 440; DB 4; Length 170;  
Best Local Similarity 90.9%; Pred. No. 8.4e-36;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60  
DB 1 LKIDSDSDYAKGERAPLOSELDTKAKLKLKLELSGKIEELDAAIXEVLKDAE 60  
QY 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99  
DB 61 GNNVEAYFKEGLEKTTAEKAEADLKKAVDEPE 99

RESULT 8  
US-08-529-055-42  
; Sequence 42, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.

```
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-42

Query Match 91.3%; Score 440; DB 4; Length 181;
Best Local Similarity 90.9%; Pred. No. 9.1e-36;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKDI DESDSE DYAK EGRA PLSQ LDTK KAK LKLE LSG KIE L DAE I XE L VQ LK DAE 60
Db 1 LK EID ESDSE DY LK EG RA PLSQ LDTK KAK LKLE LSG KIE L DAE I XE L VQ LK DAE 60

Qy 61 GNNV EAYF KEG LKTTAE KKA E LKAE A D LK KAV D EPE 99
Db 61 GNNV EAYF KEG LKTTAE KKA E LKAE A D LK KAV D EPE 99

RESULT 9
US-08-714-741-40
; Sequence 40, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue

; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-40

Query Match 91.3%; Score 440; DB 4; Length 864;
Best Local Similarity 90.9%; Pred. No. 5.7e-35;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKDI DESDSE DYAK EGRA PLSQ LDTK KAK LKLE LSG KIE L DAE I XE L VQ LK DAE 60
Db 465 LK EID ESDSE DY LK EG RA PLSQ LDTK KAK LKLE LSG KIE L DAE I XE L VQ LK DAE 524

Qy 61 GNNV EAYF KEG LKTTAE KKA E LKAE A D LK KAV D EPE 99
Db 525 GNNV EAYF KEG LKTTAE KKA E LKAE A D LK KAV D EPE 563

RESULT 10
US-08-710-749-14
; Sequence 14, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
```

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-14

Query Match 90.7%; Score 437; DB 2; Length 99;  
Best Local Similarity 89.9%; Pred. No. 8.8e-36;  
Matches 89; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60  
||:|||||  
DB 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60  
||:|||||  
QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
|||||  
DB 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
|||||

RESULT 11  
US-08-529-055-59  
; Sequence 59, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 188 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-59

Query Match 90.7%; Score 437; DB 4; Length 188;

Best Local Similarity 89.9%; Pred. No. 1.9e-35;  
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60  
||:|||||  
DB 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60  
||:|||||  
QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
|||||  
DB 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
|||||

RESULT 12  
US-08-710-749-17  
; Sequence 17, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-17

Query Match 89.2%; Score 430; DB 2; Length 99;  
Best Local Similarity 88.9%; Pred. No. 4.3e-35;  
Matches 88; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60  
||:|||||  
DB 1 LKIDSDSDYAKGERAPLQSELDTKKAKLLKLELSGKIBELDAEIXEVLQKDAE 60  
||:|||||  
QY 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
|||||  
DB 61 GNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
|||||

RESULT 13  
US-09-147-875A-10  
; Sequence 10, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:

```
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-10

Query Match      89.1%; Score 429.5; DB 4; Length 100;
Best Local Similarity 90.0%; Pred. No. 4.8e-35;
Matches 90; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 1 LKIDSDSDSYAKEGRAPLQSLDTKAKLLKLELSKGIKIELDAEIXELE-VQLKDA 59
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 LKEIDSDSDSYAKEGRAPLQSLDKLDAKAKLSKLELSKIDELDAEIAKLECVQLKDA 60

Qy 60 EGNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDE 99
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 EGNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDE 100

RESULT 14
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-2

Query Match      88.8%; Score 428; DB 4; Length 141;
Best Local Similarity 90.7%; Pred. No. 1e-34;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKIDSDSDSYAKEGRAPLQSLDTKAKLLKLELSKGIKIELDAEIXELEVLKDAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 45 LKEIDSDSDSYLKEGLRAPLQSLDKLDAKAKLSKLELSKIDELDAEIAKLEVLKDAE 104

Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 105 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDE 141

RESULT 15
US-08-714-741-42
; Sequence 42, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Votter, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
```

```
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESS: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Prommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-42

Query Match      88.8%; Score 428; DB 4; Length 588;
Best Local Similarity 90.7%; Pred. No. 5.4e-34;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKIDSDSDSYAKEGRAPLQSLDTKAKLLKLELSKGIKIELDAEIXELEVLKDAE 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 492 LKEIDSDSDSYLKEGLRAPLQSLDKLDAKAKLSKLELSKIDELDAEIAKLEVLKDAE 551

Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 552 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDE 588

Search completed: June 21, 2005, 10:25:20
Job time : 18.4867 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKIDSESDYAKGERAP.....KKAELKADLKKAUDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 480   | 99.6        | 99     | 15 | US-10-674-755-14  |
| 2          | 472   | 97.9        | 206    | 15 | US-10-299-636-69  |
| 3          | 451   | 93.6        | 99     | 15 | US-10-674-755-15  |
| 4          | 448   | 92.9        | 99     | 15 | US-10-674-755-16  |
| 5          | 440   | 91.3        | 170    | 15 | US-10-299-636-75  |
| 6          | 440   | 91.3        | 181    | 15 | US-10-299-636-57  |
| 7          | 440   | 91.3        | 643    | 15 | US-10-299-636-95  |
| 8          | 440   | 91.3        | 670    | 9  | US-09-748-875-63  |
| 9          | 440   | 91.3        | 670    | 10 | US-09-298-523B-63 |
| 10         | 440   | 91.3        | 690    | 9  | US-09-748-875-61  |
| 11         | 440   | 91.3        | 690    | 10 | US-09-298-523B-61 |

|    |       |      |     |    |                      |                   |
|----|-------|------|-----|----|----------------------|-------------------|
| 12 | 440   | 91.3 | 691 | 9  | US-09-748-875-1      | Sequence 1, Appli |
| 13 | 440   | 91.3 | 691 | 10 | US-09-298-523B-1     | Sequence 1, Appli |
| 14 | 440   | 91.3 | 701 | 9  | US-09-748-875-62     | Sequence 62, Appl |
| 15 | 440   | 91.3 | 701 | 10 | US-09-298-523B-62    | Sequence 62, Appl |
| 16 | 440   | 91.3 | 707 | 9  | US-09-748-875-2      | Sequence 2, Appli |
| 17 | 440   | 91.3 | 707 | 10 | US-09-298-523B-2     | Sequence 2, Appli |
| 18 | 440   | 91.3 | 711 | 9  | US-09-748-875-3      | Sequence 3, Appli |
| 19 | 440   | 91.3 | 711 | 10 | US-09-298-523B-3     | Sequence 3, Appli |
| 20 | 440   | 91.3 | 739 | 17 | US-10-732-923-3294   | Sequence 3294, Ap |
| 21 | 440   | 91.3 | 929 | 9  | US-09-748-875-60     | Sequence 60, Appl |
| 22 | 440   | 91.3 | 929 | 10 | US-09-298-523B-60    | Sequence 60, Appl |
| 23 | 440   | 91.3 | 929 | 15 | US-10-299-636-94     | Sequence 94, Appl |
| 24 | 437   | 90.7 | 188 | 15 | US-10-299-636-74     | Sequence 74, Appl |
| 25 | 429.5 | 89.1 | 100 | 15 | US-10-674-755-10     | Sequence 10, Appl |
| 26 | 428   | 88.8 | 141 | 14 | US-10-254-995-2      | Sequence 14, Appl |
| 27 | 428   | 88.8 | 589 | 9  | US-09-748-875-14     | Sequence 14, Appl |
| 28 | 428   | 88.8 | 589 | 10 | US-09-298-523B-14    | Sequence 97, Appl |
| 29 | 428   | 88.8 | 589 | 15 | US-10-299-636-97     | Sequence 73, Appl |
| 30 | 426   | 88.4 | 204 | 15 | US-10-299-636-73     | Sequence 11, Appl |
| 31 | 386   | 80.1 | 99  | 15 | US-10-674-755-11     | Sequence 11, Appl |
| 32 | 386   | 80.1 | 198 | 15 | US-10-299-636-76     | Sequence 76, Appl |
| 33 | 386   | 80.1 | 204 | 15 | US-10-299-636-66     | Sequence 66, Appl |
| 34 | 386   | 80.1 | 354 | 15 | US-10-299-636-105    | Sequence 105, App |
| 35 | 386   | 80.1 | 588 | 15 | US-10-299-636-96     | Sequence 96, Appl |
| 36 | 386   | 80.1 | 619 | 10 | US-09-882-774-1      | Sequence 1, Appli |
| 37 | 386   | 80.1 | 619 | 15 | US-10-282-122A-73702 | Sequence 73702, A |
| 38 | 386   | 80.1 | 619 | 16 | US-10-414-532-72     | Sequence 12, Appl |
| 39 | 372.5 | 77.3 | 100 | 15 | US-10-674-755-12     | Sequence 62, Appl |
| 40 | 365   | 75.7 | 193 | 15 | US-10-299-636-64     | Sequence 64, Appl |
| 41 | 361   | 74.9 | 99  | 15 | US-10-674-755-17     | Sequence 17, Appl |
| 42 | 359   | 74.5 | 195 | 15 | US-10-299-636-86     | Sequence 86, Appl |
| 43 | 355.5 | 73.8 | 336 | 15 | US-10-299-636-103    | Sequence 103, App |
| 44 | 355   | 73.7 | 99  | 15 | US-10-674-755-13     | Sequence 13, Appl |
| 45 | 308.5 | 64.0 | 100 | 15 | US-10-674-755-4      | Sequence 4, Appli |

#### ALIGNMENTS

##### RESULT 1

US-10-674-755-14  
; Sequence 14, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; PRIOR FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(99)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-14

|                       |        |   |        |                       |
|-----------------------|--------|---|--------|-----------------------|
| Query Match           | 99.6%  | Score 480;  | DB 15; | Length 99;            |
| Best Local Similarity | 100.0% | Pred. No. 3.5e-33;                                    |        |                       |
| Matches               | 99;    | Conservative  | 0;     | Mismatches 0; Gaps 0; |
| QY                    | 1      | LKIDSESDYAKGERAPLQSELDTKKAKLKLKELSGKIEELDAIXELVQLKDAE | 60     |                       |
| Db                    | 1      | LKIDSESDYAKGERAPLQSELDTKKAKLKLKELSGKIEELDAIXELVQLKDAE | 60     |                       |
| QY                    | 61     | GNNVAYPKGLEKTTAKKAELEKAEADLKKAUDEPE                   | 99     |                       |

Db 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99

## RESULT 2

US-10-674-636-69  
; Sequence 69, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 69  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (50)  
; OTHER INFORMATION: Xaa at position 50 is unknown  
US-10-299-636-69

Query Match 97.9%; Score 472; DB 15; Length 206;  
Best Local Similarity 99.0%; Pred. No. 3.8e-32;  
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKIDSDSDSDYAKGERAPLQSLDTKKAKLLKLELSGKIELDAEIXEVLKDAE 60  
Db 1 LKIDSDSDSDYAKGERAPLQSLDTKKAKLLKLELSGKIELDAEIXEVLKDAE 60  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99

## RESULT 3

US-10-674-755-15  
; Sequence 15, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-15

Query Match 93.6%; Score 451; DB 15; Length 99;  
Best Local Similarity 92.9%; Pred. No. 9.8e-31;  
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKIDSDSDSDYAKGERAPLQSLDTKKAKLLKLELSGKIELDAEIXEVLKDAE 60  
Db 1 LKIDSDSDSDYAKGERAPLQSLDTKKAKLLKLELSGKIELDAEIXEVLKDAE 60  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99

## RESULT 4

US-10-674-755-16  
; Sequence 16, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-16

Query Match 92.9%; Score 448; DB 15; Length 99;  
Best Local Similarity 91.9%; Pred. No. 1.8e-30;  
Matches 91; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKIDSDSDSDYAKGERAPLQSLDTKKAKLLKLELSGKIELDAEIXEVLKDAE 60  
Db 1 LKIDSDSDSDYAKGERAPLQSLDTKKAKLLKLELSGKIELDAEIXEVLKDAE 60  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99

## RESULT 5

US-10-299-636-75  
; Sequence 75, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 75  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-75

Query Match 91.3%; Score 440; DB 15; Length 170;  
Best Local Similarity 90.9%; Pred. No. 1.5e-29;

```

Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKIDSESDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 60
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKIDSESDYKGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
US-10-299-636-57
; Sequence 57, Application US/10299636
; Publication No. US2004007847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-57

Query Match 91.3%; Score 440; DB 15; Length 181;
Best Local Similarity 90.9%; Pred. No. 1.6e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKIDSESDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 60
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LKIDSESDYKGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-10-299-636-95
; Sequence 95, Application US/10299636
; Publication No. US2004007847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055

```

```

; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-95

Query Match 91.3%; Score 440; DB 15; Length 643;
Best Local Similarity 90.9%; Pred. No. 6.7e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKIDSESDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 60
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 LKIDSESDYKGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 304
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 343
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
US-09-748-875-63
; Sequence 63, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: Briles et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-63

Query Match 91.3%; Score 440; DB 9; Length 670;
Best Local Similarity 90.9%; Pred. No. 7.1e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 LKIDSESDYAKGERAPLQSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 60
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 498 LKIDSESDYKGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 557
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 558 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 596
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
US-09-298-523B-63
; Sequence 63, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: Briles et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae

```

## US-09-298-523B-63

Query Match 91.3%; Score 440; DB 10; Length 670;  
Best Local Similarity 90.9%; Pred. No. 7.1e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 60  
Db 498 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 557  
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99  
Db 558 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 596

## RESULT 10

US-09-748-875-61  
; Sequence 61, Application US/09748875  
; Publication No. US20010016200A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/748,875  
; PRIOR FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-61

Query Match 91.3%; Score 440; DB 9; Length 690;  
Best Local Similarity 90.9%; Pred. No. 7.3e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 60  
Db 529 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 588  
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99  
Db 589 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 627

## RESULT 11

US-09-298-523B-61  
; Sequence 61, Application US/09298523B  
; Publication No. US20030059438A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/298,523B  
; CURRENT FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-298-523B-61

Query Match 91.3%; Score 440; DB 10; Length 690;  
Best Local Similarity 90.9%; Pred. No. 7.3e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 60

Db 529 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 588  
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99  
Db 589 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 627

## RESULT 12

US-09-748-875-1  
; Sequence 1, Application US/09748875  
; Publication No. US20010016200A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/748,875  
; CURRENT FILING DATE: 2000-12-26  
; PRIOR FILING DATE: 1999-04-23  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 691  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-1

Query Match 91.3%; Score 440; DB 9; Length 691;  
Best Local Similarity 90.9%; Pred. No. 7.3e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 60  
Db 530 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 589  
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99  
Db 590 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 628

## RESULT 13

US-09-298-523B-1  
; Sequence 1, Application US/09298523B  
; Publication No. US20030059438A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/298,523B  
; CURRENT FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 691  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-298-523B-1

Query Match 91.3%; Score 440; DB 10; Length 691;  
Best Local Similarity 90.9%; Pred. No. 7.3e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 60  
Db 530 LKIDESDSEDYAKEGERAPLQSELDTKKAKLLKLEELSGKIEELDAAEIXELEVQLKDAE 589  
Qy 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99  
Db 590 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 628

```

RESULT 14
US-09-748-875-62
; Sequence 62, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-62

```

```

Query Match 91.3%; Score 440; DB 9; Length 701;
Best Local Similarity 90.9%; Pred. No. 7.4e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGGERAPLOSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 60
   ||:|||||  |||  |||:|||||  |||  |||:|||||  |||  |||:|||||  |||  |||:|||||
Db 529 LKEIDSDSDYAKGGERAPLOSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 588

QY 61 GNNVVEAYFKGLEKTTAEKKALEKAEADLKKAVDEPE 99
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 589 GNNVVEAYFKGLEKTTAEKKALEKAEADLKKAVDEPE 627

```

```

RESULT 15
US-09-298-523B-62
; Sequence 62, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-62

```

```

Query Match 91.3%; Score 440; DB 10; Length 701;
Best Local Similarity 90.9%; Pred. No. 7.4e-29;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDSDSDYAKGGERAPLOSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 60
   ||:|||||  |||  |||:|||||  |||  |||:|||||  |||  |||:|||||  |||  |||:|||||
Db 529 LKEIDSDSDYAKGGERAPLOSELDTKKAKLLKLEELSGKIEELDAEIXELEVQLKDAE 588

QY 61 GNNVVEAYFKGLEKTTAEKKALEKAEADLKKAVDEPE 99
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 589 GNNVVEAYFKGLEKTTAEKKALEKAEADLKKAVDEPE 627

```

Search completed: June 21, 2005, 11:18:35  
Job time : 64.2388 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKXIDSDSDYAKGERAP.....KXAELEKADLKKAVDEPE 99

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR:\*

2: PIR:\*

3: PIR:\*

4: PIR:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID               | Description        |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1          | 386   | 80.1        | 619    | 2 A97887            | surface protein ps |
| 2          | 386   | 80.1        | 619    | 2 A41971            | surface protein ps |
| 3          | 119.5 | 24.8        | 161    | 2 S48396            | tropomyosin TPM2 - |
| 4          | 116.5 | 24.2        | 744    | 2 P95013            | pneumococcal surfa |
| 5          | 112.5 | 23.3        | 281    | 2 F75216            | hypothetical prote |
| 6          | 108.5 | 22.5        | 650    | 2 A11333            | ABC transporter (A |
| 7          | 108.5 | 22.5        | 1006   | 2 C70445            | ATPase subunit of  |
| 8          | 107   | 22.2        | 1319   | 2 A28313            | glued protein - fr |
| 9          | 104.5 | 21.7        | 852    | 2 D72230            | conserved hypothet |
| 10         | 103.5 | 21.5        | 785    | 2 T01025            | hypothetical prote |
| 11         | 103.5 | 21.5        | 1110   | 2 L51116            | NF-180 - sea lamp  |
| 12         | 102   | 21.2        | 764    | 2 T05409            | hypothetical prote |
| 13         | 101.5 | 21.1        | 279    | 2 D71453            | hypothetical prote |
| 14         | 100   | 20.7        | 1169   | 2 P15 homolog - Met |                    |
| 15         | 100   | 20.7        | 1269   | 2 F84730            | probable myosin he |
| 16         | 99    | 20.5        | 233    | 2 S70531            | bbk2.11 protein pr |
| 17         | 99    | 20.5        | 650    | 2 A41704            | ABC transporter (A |
| 18         | 99    | 20.5        | 1177   | 2 T05390            | chromosome segrega |
| 19         | 98.5  | 20.4        | 779    | 2 T05990            | hypothetical prote |
| 20         | 97    | 20.1        | 1790   | 2 S67593            | transport protein  |
| 21         | 97    | 20.1        | 2139   | 2 T18296            | myosin heavy chain |
| 22         | 96.5  | 20.0        | 384    | 2 G86287            | hypothetical prote |
| 23         | 96    | 19.9        | 1410   | 1 A57013            | early endosome ant |
| 24         | 96    | 19.9        | 1875   | 2 S38173            | myosin-like protei |
| 25         | 96    | 19.9        | 3488   | 2 T34418            | hypothetical prote |
| 26         | 95.5  | 19.8        | 229    | 2 S70532            | outer surface prot |
| 27         | 95.5  | 19.8        | 1164   | 2 T24806            | hypothetical prote |
| 28         | 95.5  | 19.8        | 1179   | 2 F21190            | probable chromosom |
| 29         | 95    | 19.7        | 876    | 2 A23767            | myosin heavy chain |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 95   | 19.7 | 880  | 2 F75103 | conserved hypothet |
| 31 | 95   | 19.7 | 1078 | 2 T18352 | protein P120 - Myc |
| 32 | 95   | 19.7 | 1475 | 2 T13318 | hypothetical prote |
| 33 | 95   | 19.7 | 1690 | 2 T13030 | microtubule bindin |
| 34 | 95   | 19.7 | 1976 | 2 A59252 | myosin heavy chain |
| 35 | 94.5 | 19.6 | 392  | 2 G95258 | secreted 45 kd pro |
| 36 | 94.5 | 19.6 | 392  | 2 B98124 | general stress pro |
| 37 | 94.5 | 19.6 | 488  | 2 F97039 | hypothetical prote |
| 38 | 94.5 | 19.6 | 635  | 2 A10625 | ABC transporter At |
| 39 | 94.5 | 19.6 | 1170 | 2 A56157 | chromosome segrega |
| 40 | 94.5 | 19.6 | 1805 | 1 A64224 | hypothetical prote |
| 41 | 94   | 19.5 | 407  | 1 EDBE03 | immediate-early pr |
| 42 | 94   | 19.5 | 1156 | 2 E69444 | chromosome segrega |
| 43 | 94   | 19.5 | 1539 | 2 T18372 | repeat organellar  |
| 44 | 93.5 | 19.4 | 629  | 2 F86351 | protein T26F17.2 [ |
| 45 | 93.5 | 19.4 | 886  | 2 H69378 | conserved hypothet |

ALIGNMENTS

RESULT 1

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C:Accession: A97887

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: A97887

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <XUR>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:glf

C:Genetics:

A:Gene: pspA

Query Match 80.1%; Score 386; DB 2; Length 619;

Best Local Similarity 81.8%; Pred. No. 1.2e-19;

Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKXIDSDSDYAKGERAPLOSLDTKKALKLEELSGKTEELDAETXELVQLKDAE 60

Db 223 LKXIDSESDYAKGERAPLOSLDKLDAKKALKLEELSDKIDELDAETAKLEDQKAAE 282

QY 61 GNNVVEAYPFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99

Db 283 ENNVVEDYFKEGLEKTTIAAKAELEKTEADLKKAVNEPE 321

RESULT 2

A41971

surface protein pspA precursor - Streptococcus pneumoniae

N:Alternate names: pneumococcal surface protein A

C:Species: Streptococcus pneumoniae

C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A41971; A60282; A33134

R:Yother, J.; Briles, D.E.

J. Bacteriol. 174, 601-609, 1992

A:Title: Structural properties and evolutionary relationships of pspA, a surface protein

A:Reference number: A41971; MUID:92105030; PMID:1729249

A:Accession: A41971

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <VOT>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:G153840; PIDN:AAA27018

A>Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:P:75636)

R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

Infect. Immun. 59, 1285-1289, 1991  
A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability  
A;Reference number: A60282; MUID:91169598; PMID:2004810

A;Accession: A60282  
A;Molecule type: protein  
A;Residues: 32-76 <TAL>  
A;Experimental source: strain JY2008  
C;Genetics:

A;Gene: pspA  
F;1-31/Domain: signal sequence #status predicted <SIG>  
F;32-619/Product: surface protein pspA #status predicted <MAT>  
F;411-430/Domain: cpl repeat homology <CP01>  
F;431-450/Domain: cpl repeat homology <CP02>  
F;451-470/Domain: cpl repeat homology <CP03>  
F;471-490/Domain: cpl repeat homology <CP04>  
F;491-510/Domain: cpl repeat homology <CP05>  
F;511-530/Domain: cpl repeat homology <CP06>  
F;531-550/Domain: cpl repeat homology <CP07>  
F;551-570/Domain: cpl repeat homology <CP08>  
F;571-591/Domain: cpl repeat homology <CP09>  
F;592-611/Domain: cpl repeat homology <CP10>

Query Match 80.1%; Score 386; DB 2; Length 619;  
Best Local Similarity 81.8%; Pred. No. 1.2e-19;  
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYAKGGRAPLQSLDTTKAKLLKLELSGKIEELDAEIXEVLQKDAE 60

Db 223 LKIEDESEDYAKGGRAPLQSKDKAKLKLKLELSGKIEELDAEIXEVLQKAAE 282

Qy 61 GNNNVYAFKGELEKTTAEKKAELKAEADLKKAVIDEPE 99

Db 283 ENNVYDFKGELEKTTAAKKAELKAELEKTEADLKKAVIDEPE 321

#### RESULT 3

S48396  
tropolymyosin TPM2 - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein YIL138C  
C;Species: Saccharomyces cerevisiae  
C;Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 09-Jul-2004  
A;Accession: S48396; A56490  
R;Churcho, C.

Submitted to the EMBL Data Library, September 1994  
A;Reference number: S48310  
A;Accession: S48396  
A;Molecule type: DNA  
A;Residues: 1-161 <CHU>  
A;Cross-references: UNIPROT:P40414; GB:Z47047; EMBL:Z38059; NID:g603997; PID:g763208; MIPS:YIL138C  
J;Drees, B.; Brown, C.; Barrall, B.G.; Bretscher, A.  
J. Cell Biol. 128, 383-392, 1995

A;Title: Tropomyosin is essential in yeast, yet the TPM1 and TPM2 products perform distinct functions  
A;Reference number: A56490; MUID:95146545; PMID:7844152  
A;Accession: A56490

A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-161 <DRE>

A;Cross-references: GB:Z47047; GB:Z38059; NID:g603997; PID:g763208  
C;Genetics:

A;Gene: SGD:TPM2  
A;Cross-references: SGD:S0001400; MIPS:YIL138C  
A;Map position: 9L  
C;Superfamily: tropomyosin TPM1  
C;Keywords: cytoskeleton

Query Match 24.8%; Score 119.5; DB 2; Length 161;  
Best Local Similarity 34.7%; Pred. No. 0.081;  
Matches 35; Conservative 22; Mismatches 37; Indels 7; Gaps 3;

Qy 1 LKIDSDSDYAKGGRAPLQSLDTTKAKLLKLELSGKIEELDAEIXEVLQKDAE 58

Db 11 LKLESESWQEKYELRSQLEQNSNTEKEN--EIKLSAKNEQLDSEVSEKLSQSDTK 68

Qy 59 --AEGNNNVYAFKGELEKTTAEKKAELKAEADLKKAVIDE 97  
Db 69 QLAEDSNLFSN--NENYTKKNQDLEQLEDESEAKLKEAMDK 108

#### RESULT 4

F95013  
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: F95013

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidorn, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I.; Nelson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: F95013

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-744 <KUR>

A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:GSPDB:G  
A;Experimental source: strain TIGR4  
C;Genetics:

A;Gene: SP0117

Query Match 24.2%; Score 116.5; DB 2; Length 744;

Best Local Similarity 29.7%; Pred. No. 0.58;

Matches 35; Conservative 24; Mismatches 32; Indels 27; Gaps 4;

Qy 4 IDSDSDYAKGGRAPLQSEL-----DTKKAKL-LKLELSGKIE 43

Db 251 LKGEAEELNAKQAEELAKKQTELEKLLDSDPEKQDDELKEAEELDKADELQNKVA 310

Qy 44 ELDAEIXEVLQKDAEAGNNNVYAFKGELEKTTAEKKAELKAEADLKKAVIDE 99

Db 311 DLEKEISNLEILLGGADPEDDTAA----LQNKLAAKKAELEKQTELEKLLDSDPE 363

#### RESULT 5

F75216  
hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
A;Accession: F75216

R;anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
A;Reference number: A75001

A;Accession: F75216  
A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-281 <KAW>

A;Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CA849103  
A;Experimental source: strain Orsay

C;Genetics:  
A;Gene: PAB2181

Query Match 23.3%; Score 112.5; DB 2; Length 281;

Best Local Similarity 29.4%; Pred. No. 0.42;

Matches 35; Conservative 25; Mismatches 32; Indels 27; Gaps 4;

Qy 4 IDSDSDYAKGGRAPLQSELDTKKAKL-----LXLELSGKIEELDAEIXE 51

Db 124 VAREEYKLLKYEK--LKQEFEEVAKTAEAELESLKAELEKAEADLKKAVIDE 181

Qy 52 LEVLQKDAEAGN-----NNVYAFKGELEKTTAEKKAELKAEADLKKAVIDE 99

Db 182 LKLEKSEVVKLMEYAKAGAELEKAELEKYEKVKRE--ELERKVSLESLNVE 238

#### RESULT 6



A11333  
ABC transporter (ATP-binding protein) homolog lmo2073 [imported] - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: A11333  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: A11333  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-650 <GLA>  
A:Cross-references: UNIPROT:Q8Y519; GB:NC\_003210; PIDN:CAD00151.1; PID:g16411543; GSPDB:  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo2073  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 22.5%; Score 108.5; DB 2; Length 650;  
Best Local Similarity 32.6%; Pred. No. 1.8; Mismatches 21; Indels 21; Gaps 4;  
Matches 36; Conservative 21; Mismatches 32; Indels 32; Gaps 4;

QY 2 KQIDSDSDSEYAKGGERA-----PLQSELDTKKAKLL-----KLEELSGKIEELDAE 49  
DB 541 KELLARLDAEDRRKKGQEVATASVRKLNQEEKEQKLLRQRKKLEEEKSNEETDEKI 600  
QY 50 XELEVQLKDAEGNNVVEAFYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
DB 601 AELELQLTNPE-----VFQDHEKALEIT-----QELDAVRADGKELMEWE 641

RESULT 7  
C70445  
ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: C70445  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98156666; PMID:9537320  
A:Accession: C70445  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1006 <AQF>  
A:Cross-references: UNIPROT:O67589; GB:AE000750; NID:g2983999; PIDN:AAC07550.1; PID:g298  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: ctpB  
C:Superfamily: endopeptidase Clp ATP-binding chain  
C:Keywords: hydrolase

Query Match 22.5%; Score 108.5; DB 2; Length 1006;  
Best Local Similarity 33.6%; Pred. No. 2.8; Mismatches 21; Indels 23; Gaps 5;  
Matches 36; Conservative 21; Mismatches 27; Indels 23; Gaps 5;

QY 1 LKQIDSDSDSEYAKGGERAPLQSELDTKKAKLLK-LEELSGKIEELDAEIXELEV 54  
DB 552 IKALEEQIIEANLKG DYKE-----AQLKIEKALEKEQBLGKVGVEAKIAELKK 604  
QY 55 QLKDAEGNNVVEAFYFKEGLEKTTAAKKAELKAELE-----KAEADLKKAVDE 97  
DB 605 KIEE-----LDEKIEAEKGDYKEAEALKIEKALEKLEKXLEQ 645

glued protein - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C:Accession: A28313  
R:Swarcop, A.; Swarcop, M.; Garen, A.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987  
A:Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued ger  
A:Reference number: A28313; MUID:87317680; PMID:2819881  
A:Accession: A28313  
A:Molecule type: DNA; mRNA  
A:Residues: 1-1319 <SWA>  
A:Cross-references: UNIPROT:PI3496  
A:Note: the authors' translation is inconsistent with the nucleotide sequence in the regi  
C:Genetics:  
A:Gene: FlyBase:Gl  
A:Cross-references: FlyBase:FBgn0001108  
A:Introns: 18/2; 479/3  
C:Keywords: cytoskeleton; glycoprotein  
F:397,590,771,888,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 22.2%; Score 107; DB 2; Length 1319;  
Best Local Similarity 32.7%; Pred. No. 4.6; Mismatches 19; Indels 22; Gaps 4;  
Matches 33; Conservative 19; Mismatches 27; Indels 22; Gaps 4;

QY 1 LKQIDSDSDSEYAKGGERAPLQSELDTKKAKLLK-----EELSGKIEELDAEIXELEV 56  
DB 429 LRDLSAHDKHDYQK-----LSKELEMKRSEVTELETKELSAKIDELEIAIVADLQEQV 482  
QY 57 KDAEGNNVVEAFYFKEGLEKTTAAKKAELKAEADLKKAVDE 97  
DB 483 DAALG-----AEWVEQLAKKQWLE-----DKVKLLEE 511

RESULT 9  
D72230  
conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: D72230  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: D72230  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-852 <ARN>  
A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3670:  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM1636  
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 21.7%; Score 104.5; DB 2; Length 852;  
Best Local Similarity 30.3%; Pred. No. 4.4; Mismatches 20; Indels 11; Gaps 2;  
Matches 27; Conservative 20; Mismatches 31; Indels 11; Gaps 2;

QY 6 ESDSDSEYAKGGERAPLQSELDTKKAKLLKLE-----LSGKIEELDAEIXELEVQLKDAEG 61  
DB 506 EKIEELHRLGYSDELLEKLEKRLKIEERHSISQKITAADVQISQIENQLKEIKG 565  
QY 62 NNVEAFYFKEGLEKTTAAKKAELKAEAD 90  
DB 566 E-----IEAKRETLKEOREMDQLKSD 587

RESULT 10  
T01025  
hypothetical protein YUP8H12R.8 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004

[illegible]

```

Query Match      20.7%; Score 100; DB 2; Length 1169;
Best Local Similarity 32.4%; Pred. No. 12;
Matches 33; Conservative 22; Mismatches 39; Indels 8; Gaps 3;

QY 1 LKIDIDESDYAKGERAPLQSELDTKKAKLKLEELSGKIEELDAAIXEVLQKDAE 60
DB 799 LKRMNIEGELKILEKAKLKNID--KGLTLVKELIPKIEELNKKVSELINKKVILE 856

QY 61 GNNNVEAYFKEGLEKTTA---EKAELEKAEADLKAVDEPE 99
DB 857 KN---ISFYKESIEKNLSILEEKRYEELAKNKLKELTEKKE 895

RESULT 15
F84730
probable myosin heavy chain [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84730
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84730
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1269 <STO>
A:Cross-references: GB:AE002093; NID:G6598493; PIDN:AAC69932.2; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32240
A:Map position: 2
  
```

```

Query Match      20.7%; Score 100; DB 2; Length 1269;
Best Local Similarity 26.3%; Pred. No. 13;
Matches 36; Conservative 24; Mismatches 39; Indels 38; Gaps 4;

QY 1 LKIDIDESDYAK-----GERAPLQSELDTKKAKLKLE-----EL 38
DB 263 IKELNEKMSNEKVEALKSSAGELAAVQELALSKRLLTEQKVSSTEALIDELTQEL 322

QY 39 SGK-----IEELDAXIXEVLQKDAEGNN---VEAYFKEGLEKTTAEKKA 82
DB 323 EQKKASESRPKBELSVLQDLDAQTGKLQAKLSQEGLNSKLAELKEKELLESLSKDQEE 382

QY 83 ELEKAEADLKAVDEPE 99
DB 383 KLTANEKLAELVLEKE 399
  
```

Search completed: June 21, 2005, 10:12:00  
 Job time : 10.9 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-14

Perfect score: 482

Sequence: 1 LKIDNESDSEYKGERAP.....KKAELEKAEADLKKAVIDEPE 99

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 464   | 96.3        | 224    | 2 Q8GNS8 | Q8GNS8 streptococc |
| 2          | 464   | 96.3        | 249    | 2 Q9L575 | Q9L575 streptococc |
| 3          | 464   | 96.3        | 426    | 2 Q9LAY5 | Q9LAY5 streptococc |
| 4          | 460   | 95.4        | 99     | 2 Q8KQK4 | Q8KQK4 streptococc |
| 5          | 451   | 93.6        | 395    | 2 Q9LAY2 | Q9LAY2 streptococc |
| 6          | 451   | 93.6        | 408    | 2 Q9LAY0 | Q9LAY0 streptococc |
| 7          | 440   | 91.3        | 739    | 2 Q9RQT4 | Q9RQT4 streptococc |
| 8          | 440   | 91.3        | 820    | 2 Q9RQT1 | Q9RQT1 streptococc |
| 9          | 440   | 91.3        | 929    | 2 Q9KK19 | Q9KK19 streptococc |
| 10         | 440   | 91.3        | 929    | 2 Q9ZAV5 | Q9ZAV5 streptococc |
| 11         | 427   | 88.6        | 437    | 2 Q9LAY4 | Q9LAY4 streptococc |
| 12         | 403.5 | 83.7        | 869    | 2 Q9KK27 | Q9KK27 streptococc |
| 13         | 386   | 80.1        | 417    | 2 Q9LAY3 | Q9LAY3 streptococc |
| 14         | 386   | 80.1        | 619    | 2 Q54972 | Q54972 streptococc |
| 15         | 386   | 80.1        | 619    | 2 Q8DR10 | Q8DR10 streptococc |
| 16         | 364   | 75.5        | 415    | 2 Q9LAY1 | Q9LAY1 streptococc |
| 17         | 306.5 | 63.6        | 393    | 2 Q9LAZ3 | Q9LAZ3 streptococc |
| 18         | 305.5 | 63.4        | 222    | 2 Q9L577 | Q9L577 streptococc |
| 19         | 305.5 | 63.4        | 225    | 2 Q9L591 | Q9L591 streptococc |
| 20         | 305.5 | 63.4        | 262    | 2 Q9L576 | Q9L576 streptococc |
| 21         | 305.5 | 63.4        | 415    | 2 Q9LAY7 | Q9LAY7 streptococc |
| 22         | 304.5 | 63.2        | 394    | 2 Q9LAY6 | Q9LAY6 streptococc |
| 23         | 304.5 | 63.2        | 395    | 2 Q9LAZ1 | Q9LAZ1 streptococc |
| 24         | 302.5 | 62.8        | 246    | 2 Q9L578 | Q9L578 streptococc |
| 25         | 301.5 | 62.6        | 255    | 2 Q9L581 | Q9L581 streptococc |
| 26         | 301.5 | 62.6        | 255    | 2 Q9L586 | Q9L586 streptococc |
| 27         | 298.5 | 61.9        | 416    | 2 Q9LAY8 | Q9LAY8 streptococc |
| 28         | 298.5 | 61.9        | 194    | 2 Q9L585 | Q9L585 streptococc |
| 29         | 292.5 | 60.7        | 218    | 2 Q6UEB2 | Q6UEB2 streptococc |
| 30         | 292.5 | 60.7        | 233    | 2 Q9L568 | Q9L568 streptococc |
| 31         | 292.5 | 60.7        | 236    | 2 Q9L569 | Q9L569 streptococc |

|    |       |      |     |          |                    |
|----|-------|------|-----|----------|--------------------|
| 32 | 292.5 | 60.7 | 243 | 2 Q9L564 | Q9L564 streptococc |
| 33 | 292.5 | 60.7 | 243 | 2 Q9L567 | Q9L567 streptococc |
| 34 | 292.5 | 60.7 | 244 | 2 Q9L565 | Q9L565 streptococc |
| 35 | 292.5 | 60.7 | 247 | 2 Q9L566 | Q9L566 streptococc |
| 36 | 292.5 | 60.7 | 249 | 2 Q9L570 | Q9L570 streptococc |
| 37 | 292.5 | 60.7 | 254 | 2 Q9L563 | Q9L563 streptococc |
| 38 | 292.5 | 60.7 | 401 | 2 Q9LAZ2 | Q9LAZ2 streptococc |
| 39 | 291.5 | 60.5 | 406 | 2 Q9LAZ0 | Q9LAZ0 streptococc |
| 40 | 290.5 | 60.3 | 340 | 2 Q8KQK5 | Q8KQK5 streptococc |
| 41 | 289.5 | 60.1 | 237 | 2 Q9L592 | Q9L592 streptococc |
| 42 | 289.5 | 60.1 | 395 | 2 Q9LAY9 | Q9LAY9 streptococc |
| 43 | 280.5 | 58.2 | 207 | 2 Q8GNS9 | Q8GNS9 streptococc |
| 44 | 189.5 | 39.3 | 653 | 2 Q34097 | Q34097 streptococc |
| 45 | 173   | 35.9 | 256 | 2 Q9L595 | Q9L595 streptococc |

#### ALIGNMENTS

##### RESULT 1

Q8GNS8 PRELIMINARY; PRT; 224 AA.  
 ID Q8GNS8  
 AC Q8GNS8;  
 DT 01-MAR-2003 (TREMELrel. 23, Created)  
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
 DE PspA (Fragment).  
 GN Name=pspA;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PN124;  
 RX MEDLINE=22411996; PubMed=12354862;  
 RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,  
 RA Lorino G., Recchia S., Pantosti A., Beall B.;  
 RT "Genotypes of invasive pneumococcal isolates recently recovered from  
 RT Italian patients."  
 RL J. Clin. Microbiol. 40:3660-3665(2002).  
 DR EMBL; AF490267; AAN37735.1; .  
 DR HSSP; P00192; IAPC.  
 DR InterPro; IPR009082; His\_kin\_homodim.  
 FT NON\_TER 1  
 FT NON\_TER 224 224  
 SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 96.3%; Score 464; DB 2; Length 224;  
 Best Local Similarity 96.0%; Pred. NO. 8.6e-24;  
 Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKIDNESDSEYKGERAPLOSELDTKKAKLLKEELSGKIEELDAEIXELEVQLKDAE 60  
 |||||  
 DB 17 LKIDNESDSEYKGEFRAPLOSELDTKKAKLLKEELSGKIEELDAEIXELEVQLKDAE 76  
 |||||  
 QY 61 GNNVNAVYFKEGLEKTTAEKAELEKAEADLKKAVIDEPE 99  
 |||||  
 DB 77 GNNVNAVYFKEGLEKTTAEKAELEKAEADLKKAVIDEPE 115  
 |||||

##### RESULT 2

Q9L575 PRELIMINARY; PRT; 249 AA.  
 ID Q9L575  
 AC Q9L575;  
 DT 01-OCT-2000 (TREMELrel. 15, Created)  
 DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE PspA (Fragment).  
 GN Name=pspA;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.

```

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF255552; AAF68105.1; -.
FT NON_TER 249
FT NON_TER 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match          96.3%; Score 464; DB 2; Length 249;
Best Local Similarity 96.0%; Pred. No. 9.5e-24;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKDI DESDSE DYAKGERAPLQSELDTKAKLLKLELSGKIEELDAEIXEVLQKDAE 60
Db 74 LKDI DESDSE DYKIEGFRAPLQSELDTKAKLLKLELSGKIEELDAEIXEVLQKDAE 133

Qy 61 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 134 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 172

RESULT 3
Q9LAY5
ID Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DLB5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071810; AAF27706.1; -.
DR HSSP: P00192; 1M6T.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426
FT NON_TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CB6634 CRC64;

Query Match          96.3%; Score 464; DB 2; Length 426;
Best Local Similarity 96.0%; Pred. No. 1.6e-23;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKDI DESDSE DYAKGERAPLQSELDTKAKLLKLELSGKIEELDAEIXEVLQKDAE 60
Db 215 LKDI DESDSE DYVKEGLRAPLQSELDTKAKLLKLELSGKIEELDAEIXEVLQKDAE 274

Qy 61 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 134 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 172

```

```

Db 275 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 313

RESULT 4
Q8KQK4
ID Q8KQK4 PRELIMINARY; PRT; 99 AA.
AC Q8KQK4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL: AY082388; AAL92493.1; -.
FT NON_TER 1
FT NON_TER 99
FT NON_TER 99
SQ SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match          95.4%; Score 460; DB 2; Length 99;
Best Local Similarity 94.9%; Pred. No. 7.4e-24;
Matches 94; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKDI DESDSE DYAKGERAPLQSELDTKAKLLKLELSGKIEELDAEIXEVLQKDAE 60
Db 1 LKDI DESDSE DYVKEGLRAPLQSELDTKAKLLKLELSGKIEELDAEIXEVLQKDAE 60

Qy 61 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNV EAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 5
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 395
FT NON_TER 395

```

```
SQ SEQUENCE 395 AA; 42963 MW; 58B6EP956BCBCC1E CRC64;
Query Match 93.6%; Score 451; DB 2; Length 395;
Best Local Similarity 92.9%; Pred. No. 1.1e-22;
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKDIQSDSDYAKGERAPLOSELDTKKAKLLKLEELSGKIEELDAETXELVQLKDAE 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 225 LBEINESDSDYAKGERAPLOSKLDKAKKLLKLEELSGKIEELDAETXELVQLKDAE 284
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 61 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 285 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 323
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 6
Q9LAYO PRELIMINARY; PRT; 408 AA.
AC Q9LAYO;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
   in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 93.6%; Score 451; DB 2; Length 408;
Best Local Similarity 92.9%; Pred. No. 1.1e-22;
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKDIQSDSDYAKGERAPLOSELDTKKAKLLKLEELSGKIEELDAETXELVQLKDAE 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 228 LBEINESDSDYAKGERAPLOSKLDKAKKLLKLEELSGKIEELDAETXELVQLKDAE 287
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 61 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 288 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 326
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 7
Q9ROT4 PRELIMINARY; PRT; 739 AA.
AC Q9ROT4;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DE Hypothetical protein pspC (fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
   protein, PspC, which elicits cross-reactive antibodies to PspA and
   provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 91.3%; Score 440; DB 2; Length 820;
Best Local Similarity 90.9%; Pred. No. 1.2e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
   protein, PspC, which elicits cross-reactive antibodies to PspA and
   provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EE2F2F676ABF989 CRC64;

Query Match 91.3%; Score 440; DB 2; Length 739;
Best Local Similarity 90.9%; Pred. No. 1.1e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKDIQSDSDYAKGERAPLOSELDTKKAKLLKLEELSGKIEELDAETXELVQLKDAE 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 537 LKDIQSDSDYAKGERAPLOSKLDKAKKLLKLEELSGKIEELDAETXELVQLKDAE 596
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 61 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 597 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 635
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 8
Q9ROT1 PRELIMINARY; PRT; 820 AA.
AC Q9ROT1;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Hypothetical protein pspC (fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
   protein, PspC, which elicits cross-reactive antibodies to PspA and
   provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 91.3%; Score 440; DB 2; Length 820;
Best Local Similarity 90.9%; Pred. No. 1.2e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 LKIDESDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDAAIXEVLQKDAE 60
Db 530 LKIDESDSDYLKEGLRAPLQSKLDTKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 589
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 590 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 628

RESULT 9
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ef10;
RX MEDLINE=2188621; Pubmed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSP; P06653; IH8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 91.3%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 1.3e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDESDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDAAIXEVLQKDAE 60
Db 530 LKIDESDSDYLKEGLRAPLQSKLDTKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 589
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 590 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 628

RESULT 10
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20038319; Pubmed=10569772;
```

```
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; IH8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 91.3%; Score 440; DB 2; Length 929;
Best Local Similarity 90.3%; Pred. No. 1.3e-21;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKIDESDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDAAIXEVLQKDAE 60
Db 530 LKIDESDSDYLKEGLRAPLQSKLDTKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 589
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 590 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 628

RESULT 11
Q9LAY4 ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EI34;
RX MEDLINE=20448953; Pubmed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW binding_1; 1.
DR NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 88.6%; Score 427; DB 2; Length 437;
Best Local Similarity 88.9%; Pred. No. 4.9e-21;
Matches 88; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LKIDESDSDYAKGERAPLQSELDTKKAKLKLKLELSGKIBELDAAIXEVLQKDAE 60
Db 235 LKIDESDSDYKXGURAPLQSKLDTKAKLSKLELSDKIDELDAEIAKHVVQLKDAE 294
QY 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 99
Db 295 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKAVDEPE 333

RESULT 12
```



```
Q9KK27
ID Q9KK27 PRELIMINARY; PRT; 869 AA.
AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae."
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; Rich.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
DR SEQUENCE 869 AA; 98732 MW; AFP2B504347E0220 CRC64;

Query Match 83.7%; Score 403.5; DB 2; Length 869;
Best Local Similarity 85.9%; Pred. No. 3.5e-19;
Matches 85; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 LKIDSESDYAKGERAPLOSGLDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60
DB 537 LKEIDSESDYAKGERAPLOSGLDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 595
QY 61 GNNVYAEYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 596 GNNVYAEYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 634

RESULT 13
Q9LAY3
ID Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PepA; mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae."
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSSP; P00192; 256B.
DR NON TER 417
FT SEQUENCE 417 AA; 46960 MW; 876EAD3276506BEC CRC64;

Query Match 80.1%; Score 386; DB 2; Length 417;

Best Local Similarity 81.8%; Pred. No. 2.6e-18;
Matches 81; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 LKIDSESDYAKGERAPLOSGLDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60
DB 213 LKEIDSESDYAKGERAPLOSGLDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 272
QY 61 GNNVYAEYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 273 ENNVYDYFKEGLEKTTAEKKAELKAEADLKKAVNEPE 311

RESULT 14
Q54972
ID Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PepA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis."
RL J. Bacteriol. 174:601-609(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Signal.
FT SIGNAL 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BD840C2841CA CRC64;

Query Match 80.1%; Score 386; DB 2; Length 619;
Best Local Similarity 81.8%; Pred. No. 3.8e-18;
Matches 81; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 LKIDSESDYAKGERAPLOSGLDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 60
DB 223 LKEIDSESDYAKGERAPLOSGLDTKKAKLLKLEELSGKIEELDAEIXEVLQKDAE 282
QY 61 GNNVYAEYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 283 ENNVYDYFKEGLEKTTAEKKAELKAEADLKKAVNEPE 321

RESULT 15
Q8DR10
ID Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name=pspA; Ordered locus names=spr0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
```

Search completed: June 21, 2005, 10:22:11  
Job time : 61.3194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-15

Perfect score: 484

Sequence: 1 LSEINSESDYAKGFRAP.....KXAELEKARADLKAVDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: Genesecp16Dec04:\*

2: Genesecp1980s:\*

3: Genesecp1990s:\*

4: Genesecp2000s:\*

5: Genesecp2001s:\*

6: Genesecp2002s:\*

7: Genesecp2003as:\*

8: Genesecp2003bs:\*

9: Genesecp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID      | Description         |
|------------|-------|-------------|--------|------------|---------------------|
| 1          | 454   | 93.8        | 206    | 2 AAW14574 | Aaw14574 Streptococ |
| 2          | 454   | 93.8        | 206    | 7 ABW02608 | Abw02608 Db15c pne  |
| 3          | 454   | 93.8        | 8991   | 6 ABU08487 | Abu08487 S. pneumo  |
| 4          | 440   | 90.9        | 170    | 7 ABW02614 | Abw02614 Rct135c p  |
| 5          | 440   | 90.9        | 181    | 7 ABW02596 | Abw02596 0922134c   |
| 6          | 440   | 90.9        | 865    | 6 ABU08489 | Abu08489 S. pneumo  |
| 7          | 440   | 90.9        | 929    | 2 AAW14593 | Aaw14593 Streptococ |
| 8          | 440   | 90.9        | 929    | 2 AAY43384 | Aay43384 S. pneumo  |
| 9          | 437   | 90.3        | 188    | 2 AAW14580 | Aaw14580 Streptococ |
| 10         | 437   | 90.3        | 188    | 7 ABW02613 | Abw02613 Rct129c p  |
| 11         | 429   | 88.6        | 1231   | 6 ABU08490 | Abu08490 Fragment   |
| 12         | 428   | 88.4        | 588    | 6 ABU08491 | Abu08491 Coiled co  |
| 13         | 428   | 88.4        | 589    | 2 AAY43392 | Aay43392 PspC alph  |
| 14         | 426   | 88.0        | 204    | 2 AAW14578 | Aaw14578 Streptococ |
| 15         | 426   | 88.0        | 204    | 7 ABW02612 | Abw02612 Rct123c p  |
| 16         | 425.5 | 87.9        | 180    | 2 AAW14562 | Aaw14562 Streptococ |
| 17         | 422.5 | 87.3        | 187    | 2 AAW14579 | Aaw14579 Streptococ |
| 18         | 404   | 83.5        | 198    | 2 AAW14581 | Aaw14581 Streptococ |
| 19         | 401   | 82.9        | 198    | 7 ABW02615 | Abw02615 Rxic pneu  |
| 20         | 401   | 82.9        | 315    | 2 AAY04375 | Aay04375 Streptococ |
| 21         | 401   | 82.9        | 619    | 2 AAR63437 | Aar63437 Pneumococ  |
| 22         | 401   | 82.9        | 619    | 2 AAR87598 | Aar87598 Streptococ |
| 23         | 401   | 82.9        | 619    | 2 AAR86911 | Aar86911 Pneumococ  |
| 24         | 401   | 82.9        | 619    | 2 AAY1838  | Aay1838 Streptococ  |
| 25         | 401   | 82.9        | 619    | 5 AAE18782 | Aae18782 S. pneumo  |

|    |       |      |     |            |                     |
|----|-------|------|-----|------------|---------------------|
| 26 | 401   | 82.9 | 619 | 6 ABU45778 | Abu45778 Protein e  |
| 27 | 401   | 82.9 | 619 | 8 AD052126 | Ad052126 Streptococ |
| 28 | 401   | 82.9 | 648 | 2 AAW70336 | Aaw70336 Streptococ |
| 29 | 401   | 82.9 | 648 | 2 AAW62274 | Aaw62274 Streptococ |
| 30 | 401   | 82.9 | 648 | 2 AAY41837 | Aay41837 Streptococ |
| 31 | 401   | 82.9 | 648 | 2 AAW87879 | Aaw87879 A. pneumoc |
| 32 | 401   | 82.9 | 653 | 2 AAW92456 | Aaw92456 S. pneumo  |
| 33 | 401   | 82.9 | 684 | 2 AAR73912 | Aar73912 Streptococ |
| 34 | 393   | 81.2 | 204 | 2 AAW14571 | Aaw14571 Streptococ |
| 35 | 393   | 81.2 | 204 | 7 ABW02605 | Abw02605 Ef1019c p  |
| 36 | 381   | 78.7 | 653 | 2 AAR27150 | Aar27150 PspA frag  |
| 37 | 378.5 | 78.2 | 289 | 2 AAW62276 | Aaw62276 Streptococ |
| 38 | 378.5 | 78.2 | 289 | 2 AAY41840 | Aay41840 Streptococ |
| 39 | 378.5 | 78.2 | 289 | 2 AAW87910 | Aaw87910 Protein s  |
| 40 | 378.5 | 78.2 | 289 | 2 AAW92458 | Aaw92458 S. pneumo  |
| 41 | 377   | 77.9 | 190 | 2 AAW14569 | Aaw14569 Streptococ |
| 42 | 377   | 77.9 | 193 | 7 ABW02603 | Abw02603 Bg9163c p  |
| 43 | 374   | 77.3 | 195 | 2 AAW14591 | Aaw14591 Streptococ |
| 44 | 374   | 77.3 | 195 | 7 ABW02625 | Abw02625 Wu2c pneu  |
| 45 | 357   | 73.8 | 623 | 6 ABU08494 | Abu08494 Fragment   |

## ALIGNMENTS

RESULT 1  
AAW14574  
ID AAW14574 standard; protein; 206 AA.

XX AAW14574;

XX 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX Streptococcus pneumoniae PspA central region.

DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;

KW bacteraemia; pneumonia.

XX Streptococcus pneumoniae; strain Db15.

OS Key Location/Qualifiers

FH Misc-difference 50 /note= "unidentified amino acid"

FT W09709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SRP-1995; 95US-00529055.

PA (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used

PT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the

CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Db15.

CC Comparison of the N-terminal and central regions (AAW14533-57 and

CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can

CC be used to divide the strains into several families based on sequence

CC homologies. PspA polypeptides, or fragments of them, can be used in

CC vaccines to protect animals against S. pneumoniae infection and hence for

CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX  
 SQ Sequence 206 AA;  
 Query Match 93.8%; Score 454; DB 2; Length 206;  
 Best Local Similarity 92.9%; Pred. No. 7.3e-36;  
 Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 LEETNESDSYAKGFRAPLQSKLDAKAKLLKLEELSGKIEELDRAIEALEVQLKDAE 60  
 Db 1 LKIDIDESDSYAKGFRAPLQSELDTYKAKLLKLEELSGKIEELDRAIEALEVQLKDAE 60  
 Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDPE 99  
 Db 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDPE 99

RESULT 2  
 ABW02608  
 ID ABW02608 standard; protein; 206 AA.  
 XX  
 AC ABW02608;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX

XX Db15c pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 XX immunological; gene therapy; immunostimulant.

XX Unidentified.

XX Key Location/Qualifiers  
 FH Misc-difference 1. .206  
 FT /note= "Xaa = Unknown amino acid"  
 FT  
 FT

XX US6592876-B1.

XX 15-JUL-2003.

XX 15-SEP-1995; 95US-00529055.

XX 20-APR-1993; 93US-00048896.

XX 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.

XX Example 6; SEQ ID NO 54; 121pp; English.

XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Db15c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 206 AA;

Query Match 93.8%; Score 454; DB 7; Length 206;  
 Best Local Similarity 92.9%; Pred. No. 7.3e-36;  
 Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 LEETNESDSYAKGFRAPLQSKLDAKAKLLKLEELSGKIEELDRAIEALEVQLKDAE 60  
 Db 1 LKIDIDESDSYAKGFRAPLQSELDTYKAKLLKLEELSGKIEELDRAIEALEVQLKDAE 60  
 Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDPE 99  
 Db 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDPE 99

RESULT 3  
 ABU08487  
 ID ABU08487 standard; protein; 8991 AA.  
 XX  
 AC ABU08487;  
 XX  
 DT 24-JUN-2003 (first entry)  
 XX

XX S. pneumoniae pneumococcal surface protein A (PspA) protein.

XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.

XX Streptococcus pneumoniae.

XX Key Location/Qualifiers  
 FH Misc-difference 1. .8991  
 FT /note= "All Xaa residues within this sequence are  
 FT unknown"  
 FT

XX US6500613-B1.

XX 31-DEC-2002.

XX 16-SEP-1996; 96US-00714741.

XX 15-SEP-1995; 95US-00529055.

XX (UYAL-) UNIV ALABAMA.

XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;

XX Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 2003-361534/34.

XX Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.

XX Disclosure; Col 145-188; 186pp; English.

XX The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents S. pneumoniae  
 CC PspA protein

XX Sequence 8991 AA;

SQ

```

Db      1 LKEIDESDSEDYLYKEGLRAPLQSKLDTKKAUKLSKEELSDKIDELDAETAKLEVLQKDAE 60
Qy      61 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKAVIDEPE 99
        |||||
Db      61 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKAVIDEPE 99
        |||||

RESULT 5
ABW02596
ID      ABW02596 standard; protein; 181 AA.
XX
XX      AC      ABW02596;
XX
XX      DT      12-FEB-2004 (first entry)
XX
XX      DE      0922134C pneumococcal surface protein A (PspA) central region.
XX
XX      KW      Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX      KW      immunological; gene therapy; immunostimulant.
XX
XX      OS      Unidentified.
XX
XX      PN      US6592876-B1.
XX
XX      PD      15-JUL-2003.
XX
XX      PP      15-SEP-1995; 95US-00529055.
XX
XX      PR      20-APR-1993; 93US-00048896.
XX      PR      06-JUN-1995; 95US-00465746.
XX
XX      PA      (UABR-) UAB RES FOUND.
XX
XX      PI      Briles DB, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX      WPI; 2003-862841/80.
XX
XX      PT      Immunological composition for obtaining expression products used for
XX      PT      detecting the presence of Streptococcus pneumoniae or its strain,
XX      PT      comprises at least two different full length isolated gene encoding
XX      PT      pneumococcal surface protein A.
XX
XX      PS      Example 6; SEQ ID NO 42; 121pp; English.
XX
XX      CC      The present invention relates to an immunological composition comprising
XX      CC      at least 2 different full length isolated genes encoding pneumococcal
XX      CC      surface protein A (PspAs) from different groups based on restriction
XX      CC      fragment polymorphism analysis. The invention is useful for obtaining
XX      CC      expression products by recombinant techniques to detect, determine,
XX      CC      isolate or diagnose the presence of Streptococcus pneumoniae or its
XX      CC      strain. The expression product is useful for preparing antigenic,
XX      CC      immunological or vaccine compositions, for eliciting antibodies, an
XX      CC      immunological response (other than or additional to antibodies) or a
XX      CC      protective response (including antibody or other immunological response
XX      CC      by administering compositions to a host). The invention is also useful as
XX      CC      vaccines and in gene therapy. The present sequence is 0922134C
XX      CC      pneumococcal surface protein A (PspA) central region. This sequence is
XX      CC      used in the exemplification of the invention
XX
XX      SQ      Sequence 181 AA;

Query Match      90.9%; Score 440; DB 7; Length 181;
Best Local Similarity 90.9%; Pred. No. 1.4e-34;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy      1 LEINSDSEDYAKGFRAPLQSKLDKAKUKLKLKEELSGKTELDARTAEILSVQKDAE 60
        |||||
Db      1 LKEIDESDSEDYLYKEGLRAPLQSKLDTKKAUKLSKEELSDKIDELDAETAKLEVLQKDAE 60
        |||||
Qy      61 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKAVIDEPE 99
        |||||
Db      61 GNNVVEAYFKEGLEKTTAEKKAELEKAEADLKKAVIDEPE 99
        |||||

```

```

RESULT 6
ABU08489
ID ABU08489 standard; protein; 865 AA.
XX AC ABU08489;
XX DT 24-JUN-2003 (first entry)
XX DE S. pneumoniae pneumococcal surface protein C (PspC) protein.
XX KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
XX KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
XX KW antibacterial.
XX OS Streptococcus pneumoniae.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Signal_peptide
XX FT 38..865
XX FT /label= Mature_PspC_protein
XX US6500613-B1.
XX PN 31-DEC-2002.
XX PD 16-SEP-1996; 96US-00714741.
XX PF 15-SEP-1995; 95US-00529055.
XX PR (UYAL-) UNIV ALABAMA.
XX PA Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX DR WPI; 2003-361534/34.
XX DR N-PSDB; ABX95377.
XX PT Isolated PspC amino acid sequence used as polymerase chain reaction or
XX PT hybridization probe, comprises pneumococcal surface protein having alpha-
XX PT helical, proline rich and repeat regions.
XX PS Claim 3; Fig 21; 186pp; English.
XX CC The present invention relates to the isolation of Streptococcus
XX CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX CC like protein having alpha-helical, proline rich and repeat regions. The
XX CC PspC and PspA proteins may be used in a vaccine to protect against
XX CC pneumococcal infections. The polynucleotide sequences encoding PspC and
XX CC PspA may be used for the expression of the proteins, and as PCR primers
XX CC or hybridisation probes. The present sequence represents S. pneumoniae
XX CC PspC protein
XX SQ Sequence 865 AA;

Query Match 90.9%; Score 440; DB 6; Length 865;
Best Local Similarity 90.9%; Pred. No. 9.4e-34;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINSDSEYAKGFRAPLQSKLDAAKALKLKEELSGLIEDAEIAELEVLQKDAE 60
Db 466 LKEIDSDSEYDLKEGLRAPLQSKLDYTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 525

Qy 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99
Db 526 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 564

RESULT 7
AAW14593
ID AAW14593 standard; protein; 929 AA.

```

```

XX AC AAW14593;
XX DT 17-OCT-2003 (revised)
XX DT 27-OCT-1997 (first entry)
XX DE Streptococcus pneumoniae PspC.
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX KW bacteraemia; pneumonia.
XX OS Streptococcus pneumoniae; strain EF6796.
XX FH Key Location/Qualifiers
XX FT Peptide 1..37
XX FT Protein /label= Sig_peptide
XX FT 38..929
XX FT /label= Mat_protein
XX FT 38..637
XX FT /label= Alpha-helix
XX FT 41..242
XX FT /label= Repeat_1
XX FT /note= "alpha-helical repeat region"
XX FT 69..637
XX FT /label= Coiled-coil
XX FT /note= "breaks in 7-residue periodicity of the coiled-
XX FT coil occur at amino acids 136-141, 261-304 and 383-387"
XX FT 332..492
XX FT /label= Repeat_2
XX FT /note= "alpha-helical repeat region"
XX FT 627..689
XX FT /label= Proline-rich
XX FT 913..929
XX FT /label= C-terminal
XX PN W09709994-A1.
XX PD 20-MAR-1997.
XX PF 16-SEP-1996; 96WO-US014819.
XX PR 15-SEP-1995; 95US-00529055.
XX PA (UABR-) UAB RES FOUND.
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX DR WPI; 1997-202002/18.
XX DR N-PSDB; AAT61728.
XX PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX PT in vaccines for protecting animals against S.pneumoniae infection.
XX PS Disclosure; Fig 13; 296pp; English.
XX CC This sequence comprises the pneumococcal protein surface C (pspC) of
XX CC Streptococcus pneumoniae strain EF6796. The sequence was deduced from the
XX CC pspC gene (AAT61728). Like PspA, PspC has an alpha-helical coiled-coil
XX CC region, proline-rich central region, repeat regions, with a choline
XX CC binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides
XX CC share 3 regions of high sequence identity. One is a protection-eliciting
XX CC region present within the alpha-helical domain. The others are the
XX CC proline-rich domain and a repeat domain shared with other choline-binding
XX CC proteins and thought to play a role in cell surface association. PspC and
XX CC PspA polypeptides, and their fragments, can be used in vaccines to
XX CC protect against S. pneumoniae infection and hence for the prevention of
XX CC diseases such as otitis media, meningitis, bacteraemia and pneumonia.
XX CC (Updated on 17-OCT-2003 to standardise OS field)
XX SQ Sequence 929 AA;

Query Match 90.9%; Score 440; DB 2; Length 929;

```

Best Local Similarity 90.9%; Pred. No. 1e-33; Mismatches 4; Indels 0; Gaps 0;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LBEINESDSEYAKGFRAPLOSKLDKAKKLLKLELSGKIELDAETAELEVLKDAE 60  
ID AAY43384 standard; protein; 929 AA.  
DB 530 LKLEIDESDSEYAKGFRAPLOSKLDKAKKLLKLELSGKIELDAETAELEVLKDAE 589

QY 61 GNNVYAYFKEGKLEKTTAEKAELEKAEADLKKAVIDEPE 99  
DB 590 GNNVYAYFKEGKLEKTTAEKAELEKAEADLKKAVIDEPE 628

RESULT 8  
AAY43384  
XX ID AAY43384 standard; protein; 929 AA.  
XX AC AAY43384;  
XX DT 27-JAN-2000 (first entry)  
XX DE S. pneumoniae PspC protein sequence.  
XX KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;  
XX KM epitopic region; immunogenic composition; vaccine composition; therapy;  
XX KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;  
XX KW anti-PspA antibody.  
XX OS Streptococcus pneumoniae.  
XX PN WO9953940-A1.  
XX PD 28-OCT-1999.  
XX PF 23-APR-1999; 99WO-US008895.  
XX PR 23-APR-1998; 98US-0082728P.  
XX PA (UYAL-) UNIV ALABAMA.  
XX PI Briles DE, Hollingshead SK, Brooks-Walter A;  
XX WPI; 1999-620581/53.  
XX DR N-PSDB; AAZ31956.  
XX CC New epitope polypeptides of pneumococcal surface protein C, used to  
PT develop products for immunological, immunogenic or vaccine compositions,  
PT particularly against Streptococcus pneumoniae infections.  
XX PS Example; Fig 11; 109pp; English.  
XX CC This sequence is the Streptococcus pneumoniae pneumococcal surface  
CC protein C. The invention relates to an isolated and/or purified  
CC polypeptide comprising at least one epitope or epitopic region of  
CC pneumococcal surface protein C (PspC). The polypeptides or vectors  
CC containing sequence encoding them can be used as immunogenic,  
CC immunological or vaccine compositions. The compositions can be used for  
CC eliciting an immunological response against Streptococcus pneumoniae  
CC (sp), which can cause otitis media, meningitis, bacteraemia and  
CC pneumonia. They can be used for eliciting an anti-PspA antibody. The  
CC nucleic acid molecules can also be used for detecting pspC, pspA or SP  
SQ Sequence 929 AA;

Query Match 90.9%; Score 440; DB 2; Length 929;  
Best Local Similarity 90.9%; Pred. No. 1e-33; Mismatches 4; Indels 0; Gaps 0;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LBEINESDSEYAKGFRAPLOSKLDKAKKLLKLELSGKIELDAETAELEVLKDAE 60  
DB 530 LKLEIDESDSEYAKGFRAPLOSKLDKAKKLLKLELSGKIELDAETAELEVLKDAE 589

QY 61 GNNVYAYFKEGKLEKTTAEKAELEKAEADLKKAVIDEPE 99  
DB 590 GNNVYAYFKEGKLEKTTAEKAELEKAEADLKKAVIDEPE 628

Db 590 GNNVYAYFKEGKLEKTTAEKAELEKAEADLKKAVIDEPE 628

RESULT 9  
AAW14580  
ID AAW14580 standard; protein; 188 AA.  
XX AC AAW14580;  
XX DT 17-OCT-2003 (revised)  
XX DT 28-OCT-1997 (first entry)  
XX DE Streptococcus pneumoniae PspA central region.  
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
XX KW bacteraemia; pneumonia.  
XX OS Streptococcus pneumoniae; strain Rct135.  
XX PN WO9709994-A1.  
XX PD 20-MAR-1997.  
XX PF 16-SEP-1996; 96WO-US014819.  
XX PR 15-SEP-1995; 95US-00529055.  
XX PA (UABR-) UAB RES FOUND.  
XX PI Briles DE, Medaniel LS, Swiatlo E, Yother J, Crain MJ;  
XX PI Hollingshead S, Tart R, Brooks-Walter A;  
XX WPI; 1997-202002/18.  
XX CC Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
PT in vaccines for protecting animals against S.pneumoniae infection.  
XX PS Example 6; Fig 13; 296pp; English.  
XX CC This sequence shows the central portion, including the C-terminus of the  
CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
XX SQ Sequence 188 AA;

Query Match 90.3%; Score 437; DB 2; Length 188;  
Best Local Similarity 89.9%; Pred. No. 2.9e-34;  
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LBEINESDSEYAKGFRAPLOSKLDKAKKLLKLELSGKIELDAETAELEVLKDAE 60  
DB 1 LKLEIDESDSEYAKGFRAPLOSKLDKAKKLLKLELSGKIELDAETAELEVLKDAE 60

QY 61 GNNVYAYFKEGKLEKTTAEKAELEKAEADLKKAVIDEPE 99  
DB 61 GNNVYAYFKEGKLEKTTAEKAELEKAEADLKKAVIDEPE 99

RESULT 10  
ABW02613  
ID ABW02613 standard; protein; 188 AA.  
XX AC ABW02613;

XX 12-FEB-2004 (first entry)  
 XX Rct129c pneumococcal surface protein A (PspA) central region.  
 DE  
 XX  
 XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 XX Unidentified.  
 OS  
 XX US6592876-B1.  
 PN  
 XX 15-JUL-2003.  
 PD  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 PF  
 XX 20-APR-1993; 93US-00048896.  
 PR  
 XX 06-JUN-1995; 95US-00465746.  
 PR  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 PI WPI; 2003-862841/80.  
 XX  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 XX Example 6; SEQ ID NO 59; 121pp; English.  
 PS  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Rct129c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 XX Sequence 188 AA;  
 SQ  
 Query Match 90.3%; Score 437; DB 7; Length 188;  
 Best Local Similarity 89.9%; Pred. No. 2.9e-34;  
 Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLELSKGIEELDAEIAEVLQKDAE 60  
 Db 1 LKEIDESESDYLKGLRPLQSKLDKAKKLLKLELSKGIEELDAEIAEVLQKDAE 60  
 Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
 Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
 RESULT 11  
 ABU08490  
 ID ABU08490 standard; protein; 1231 AA.  
 XX  
 XX ABU08490;  
 AC  
 XX 24-JUN-2003 (first entry)  
 DT  
 XX Fragment of S. pneumoniae PspC protein.  
 DE  
 XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;

KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 XX antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN US6500613-B1.  
 XX  
 XX 31-DEC-2002.  
 PD  
 XX 16-SEP-1996; 96US-00714741.  
 PF  
 XX 15-SEP-1995; 95US-00529055.  
 PR  
 XX (UYAL-) UNIV ALABAMA.  
 XX  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A,  
 XX WPI; 2003-361534/34.  
 DR  
 XX Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.  
 PT  
 XX Disclosure; Fig 22; 186pp; English.  
 PS  
 XX The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents a fragment of S.  
 CC pneumoniae PspC protein  
 XX  
 XX Sequence 1231 AA;  
 SQ  
 Query Match 88.6%; Score 429; DB 6; Length 1231;  
 Best Local Similarity 90.7%; Pred. No. 1.7e-32;  
 Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 Qy 3 EINESDSEYAKGFRAPLQSKLDKAKKLLKLELSKGIEELDAEIAEVLQKDAEGN 62  
 Db 494 EVQLSESDYAKGFRAPLQSKLDKAKKLLKLELSKGIEELDAEIAEVLQKDAEGN 553  
 Qy 63 NNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
 Db 554 NNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 590  
 RESULT 12  
 ABU08491  
 ID ABU08491 standard; protein; 588 AA.  
 XX  
 XX AC ABU08491;  
 AC  
 XX 24-JUN-2003 (first entry)  
 DT  
 XX Coiled coil motif of alpha-helix of S. pneumoniae PspC protein.  
 DE  
 XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 KW  
 XX Streptococcus pneumoniae.  
 OS  
 XX US6500613-B1.  
 PN  
 XX 31-DEC-2002.  
 PD  
 XX 16-SEP-1996; 96US-00714741.  
 PF  
 XX





CC alpha-helix region and some of the proline-rich region, of pneumococcal  
CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.  
CC Comparison of the N-terminal and central regions (AAW14533-57 and  
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
CC be used to divide the strains into several families based on sequence  
CC homologies. PspA polypeptides, or fragments of them, can be used in  
CC vaccines to protect animals against S. pneumoniae infection and hence for  
CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
CC region and the immediate 5' tip of the coding sequence are likely to be  
CC the critical sequences for predicting PspA cross-reactions and vaccine  
CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX  
XX  
SQ Sequence 204 AA;  
Query Match 88.0%; Score 426; DB 2; Length 204;  
Best Local Similarity 87.9%; Pred. No. 3.7e-33;  
Matches 87; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 LEEINESDSYAKGFRAPLQSKLDAAKAKLLKLELSKGIELDAAIEAEVQLKDAE 60  
Db 1 IKEXDESXSDYLKEGLRAPLQSKLDTKAKLSKLELSDKIDELDAEIAKLEVLKDAE 60  
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 15  
ABW02612  
ID ABW02612 standard; protein; 204 AA.  
XX  
XX AC ABW02612;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
XX Rct123c pneumococcal surface protein A (PspA) central region.  
XX  
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
XX immunological; gene therapy; immunostimulant.  
XX  
XX Unidentified.

Key Location/Qualifiers  
FH Misc-difference 1. .204  
FT /note= "Xaa = Unknown amino acid"  
XX  
XX US6592876-B1.  
XX  
XX 15-JUL-2003.  
XX  
XX 15-SEP-1995; 95US-00529055.  
XX  
XX 20-APR-1993; 93US-00048896.  
XX  
XX 06-JUN-1995; 95US-00465746.  
XX  
XX (UABR-) UAB RES FOUND.  
XX  
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX  
XX WPI; 2003-862841/80.

XX  
XX Immunological composition for obtaining expression products used for  
XX detecting the presence of Streptococcus pneumoniae or its strain,  
XX comprises at least two different full length isolated gene encoding  
XX pneumococcal surface protein A.  
XX  
XX Example 6; SEQ ID NO 58; 121bp; English.  
XX  
XX The present invention relates to an immunological composition comprising  
XX at least 2 different full length isolated genes encoding pneumococcal  
XX surface protein A (PspAs) from different groups based on restriction  
XX fragment polymorphism analysis. The invention is useful for obtaining

CC expression products by recombinant techniques to detect, determine,  
CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
CC strain. The expression product is useful for preparing antigenic,  
CC immunological or vaccine compositions, for eliciting antibodies, an  
CC immunological response (other than or additional to antibodies) or a  
CC protective response (including antibody or other immunological response  
CC by administering compositions to a host). The invention is also useful as  
CC vaccines and in gene therapy. The present sequence is Rct123c  
CC pneumococcal surface protein A (PspA) central region. This sequence is  
XX used in the exemplification of the invention  
XX  
SQ Sequence 204 AA;

Query Match 88.0%; Score 426; DB 7; Length 204;  
Best Local Similarity 87.9%; Pred. No. 3.7e-33;  
Matches 87; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 LEEINESDSYAKGFRAPLQSKLDAAKAKLLKLELSKGIELDAAIEAEVQLKDAE 60  
Db 1 IKEXDESXSDYLKEGLRAPLQSKLDTKAKLSKLELSDKIDELDAEIAKLEVLKDAE 60  
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

Search completed: June 21, 2005, 10:10:15  
Job time : 74.8459 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-15

Perfect score: 484

Sequence: 1 LBEINESSEDYAKGFRAP.....KKAELKAEADLKKAVDPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata1/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/prodata1/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/prodata1/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/prodata1/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/prodata1/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/prodata1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 484   | 100.0       | 99     | 4     | US-09-147-875A-15 |
| 2          | 470   | 97.1        | 99     | 2     | US-08-710-749-14  |
| 3          | 454   | 93.8        | 206    | 4     | US-08-529-055-54  |
| 4          | 454   | 93.8        | 8991   | 4     | US-08-714-741-32  |
| 5          | 451   | 93.2        | 99     | 4     | US-09-147-875A-14 |
| 6          | 445   | 91.9        | 99     | 2     | US-08-710-749-17  |
| 7          | 444.5 | 91.8        | 100    | 4     | US-09-147-875A-10 |
| 8          | 440   | 90.9        | 99     | 2     | US-08-710-749-13  |
| 9          | 440   | 90.9        | 170    | 4     | US-08-529-055-60  |
| 10         | 440   | 90.9        | 181    | 4     | US-08-529-055-42  |
| 11         | 440   | 90.9        | 864    | 4     | US-08-714-741-40  |
| 12         | 437   | 90.3        | 99     | 4     | US-09-147-875A-16 |
| 13         | 437   | 90.3        | 188    | 4     | US-08-529-055-59  |
| 14         | 429   | 88.6        | 1231   | 4     | US-08-714-741-41  |
| 15         | 428   | 88.4        | 141    | 4     | US-09-286-981B-2  |
| 16         | 428   | 88.4        | 588    | 4     | US-08-714-741-42  |
| 17         | 426   | 88.0        | 99     | 2     | US-08-710-749-15  |
| 18         | 426   | 88.0        | 204    | 4     | US-08-529-055-58  |
| 19         | 401   | 82.9        | 99     | 2     | US-08-710-749-11  |
| 20         | 401   | 82.9        | 198    | 4     | US-08-529-055-61  |
| 21         | 401   | 82.9        | 619    | 1     | US-08-465-746-2   |
| 22         | 401   | 82.9        | 619    | 1     | US-08-214-164-2   |
| 23         | 401   | 82.9        | 619    | 2     | US-08-467-852A-3  |
| 24         | 401   | 82.9        | 619    | 2     | US-08-246-636-2   |
| 25         | 401   | 82.9        | 619    | 2     | US-08-247-491A-3  |
| 26         | 401   | 82.9        | 619    | 2     | US-08-319-795-2   |
| 27         | 401   | 82.9        | 619    | 2     | US-08-468-985-2   |

|    |       |      |     |   |                   |                   |
|----|-------|------|-----|---|-------------------|-------------------|
| 28 | 401   | 82.9 | 619 | 3 | US-08-312-949-2   | Sequence 2, Appli |
| 29 | 401   | 82.9 | 648 | 1 | US-08-072-070-2   | Sequence 2, Appli |
| 30 | 401   | 82.9 | 648 | 1 | US-08-469-434-2   | Sequence 2, Appli |
| 31 | 401   | 82.9 | 648 | 1 | US-08-214-222-2   | Sequence 2, Appli |
| 32 | 401   | 82.9 | 648 | 2 | US-08-467-852A-2  | Sequence 2, Appli |
| 33 | 401   | 82.9 | 648 | 2 | US-08-468-718-2   | Sequence 2, Appli |
| 34 | 401   | 82.9 | 648 | 2 | US-08-247-491A-2  | Sequence 2, Appli |
| 35 | 401   | 82.9 | 648 | 3 | US-08-446-201-3   | Sequence 3, Appli |
| 36 | 401   | 82.9 | 695 | 1 | US-08-127-499A-23 | Sequence 23, Appl |
| 37 | 401   | 82.9 | 695 | 1 | US-08-482-847-23  | Sequence 23, Appl |
| 38 | 393   | 81.2 | 99  | 2 | US-08-710-749-10  | Sequence 10, Appl |
| 39 | 393   | 81.2 | 99  | 4 | US-09-147-875A-11 | Sequence 11, Appl |
| 40 | 393   | 81.2 | 204 | 4 | US-08-529-055-51  | Sequence 51, Appl |
| 41 | 389   | 80.4 | 288 | 3 | US-08-312-949-4   | Sequence 4, Appli |
| 42 | 389   | 80.4 | 288 | 3 | US-08-446-201-4   | Sequence 4, Appli |
| 43 | 387.5 | 80.1 | 100 | 4 | US-09-147-875A-12 | Sequence 12, Appl |
| 44 | 378.5 | 78.2 | 289 | 1 | US-08-072-070-4   | Sequence 4, Appli |
| 45 | 378.5 | 78.2 | 289 | 1 | US-08-469-434-4   | Sequence 4, Appli |

#### ALIGNMENTS

##### RESULT 1

US-09-147-875A-15  
; Sequence 15, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-15

Query Match 100.0%; Score 484; DB 4; Length 99;  
Best Local Similarity 100.0%; Pred. No. 1.9e-40;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |    |  |    |
|----|----|--|----|
| QY | 1  | LBEINESSEDYAKGFRAPLOSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE | 60 |
| DB | 1  | LBEINESSEDYAKGFRAPLOSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE | 60 |
| QY | 61 | GNNVEAYPKGLEKTTAEKKAELKAEADLKKAVDPE                      | 99 |
| DB | 61 | GNNVEAYPKGLEKTTAEKKAELKAEADLKKAVDPE                      | 99 |

##### RESULT 2

US-08-710-749-14  
; Sequence 14, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA: US/08/710,749  
;; FILING DATE: 20-SEP-1996  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2074  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 14:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 99 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: n/a  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: amino acid  
;; US-08-710-749-14

Query Match 97.1%; Score 470; DB 2; Length 99;  
Best Local Similarity 97.0%; Pred. No. 4.5e-39;  
Matches 96; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LEEINESDSEDYAKGPRAPLQSKLDAKAKLKLKLELSGKIBELDAEIAELEVLKDAE 60  
Db 1 LEEINESDSEDYAKGPRAPLQSKLDAKAKLKLKLELSGKIBELDAEIAELEVLKDXVE 60  
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

## RESULT 3

US-08-529-055-54  
; Sequence 54, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/529,055  
;; FILING DATE: 15-SEP-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer, William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2400  
;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 54:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 206 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-529-055-54

Query Match 93.8%; Score 454; DB 4; Length 206;  
Best Local Similarity 92.9%; Pred. No. 4.1e-37;  
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LEEINESDSEDYAKGPRAPLQSKLDAKAKLKLKLELSGKIBELDAEIAELEVLKDAE 60  
Db 1 LKIDIESDSEDYAKGPRAPLQSELDTTKAKLKLKLELSGKIBELDAEIAELEVLKDAE 60  
Qy 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

## RESULT 4

US-08-714-741-32  
; Sequence 32, Application US/08714741  
; Patent No. 6500613  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
; TITLE OF INVENTION: PORTIONS AND PRODUCTS  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.  
; ZIP: 10036

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/714,741  
;; FILING DATE: 16-SEP-1996  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Frommer Esq., William S.  
;; REGISTRATION NUMBER: 25,506  
;; REFERENCE/DOCKET NUMBER: 454312-2460  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 840-3333  
;; TELEFAX: (212) 840-0712  
;; INFORMATION FOR SEQ ID NO: 32:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 8991 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: amino acid  
;; US-08-714-741-32

```

US-09-147-8/3A-10
Query Match          91.8%; Score 444.5; DB 4; Length 100;
Best Local Similarity 93.0%; Pred. No. 1.5e-36;
Matches 93; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Qy 1 LERINESDSDYAKGFRAPLQSKLDAAKAKLLKLELSQKIELDRAIELE-VQLKDA 59
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1 LKEIDESDSDYAKGFRAPLQSKLDAAKAKLLKLELSQKIELDRAIEIAKLECVQLKDA 60
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 60 EGNNV EAYFKEGLEKTTAEKKA ELEKAEADLKKA VDEPE 99
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 61 EGNNV EAYFKEGLEKTTAEKKA ELEKAEADLKKA VDEPE 100
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 8
US-08-710-749-13
; Sequence 13, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28

```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-13

Query Match 90.9%; Score 440; DB 2; Length 99;
Best Local Similarity 89.9%; Pred. No. 4e-36;
Matches 89; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LEEINESDSYAKGFRAPLQSKLDKAKKLLKLELSGKIELDAEIAELEVLQKDAE 60
Db 1 LKIDESDSYAKGFRAPLQSELDTKKAKLLKLELSGKIELDAEIAELEVLQKDE 60

Qy 61 GNNVYAFYFKEGLKTTAAKAELEKAEADLKKAVDPE 99
Db 61 GNNVYAFYFKEGLKTTAAKATELEKAEADLKKAVDPE 99

RESULT 9
US-08-529-055-60
; Sequence 60, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-13
```

```
;
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-60

Query Match 90.9%; Score 440; DB 4; Length 170;
Best Local Similarity 90.9%; Pred. No. 7.7e-36;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSYAKGFRAPLQSKLDKAKKLLKLELSGKIELDAEIAELEVLQKDAE 60
Db 1 LKIDESDSYAKGFRAPLQSKLDTKKAKLLKLELSGKIELDAEIAELEVLQKDAE 60

Qy 61 GNNVYAFYFKEGLKTTAAKAELEKAEADLKKAVDPE 99
Db 61 GNNVYAFYFKEGLKTTAAKAELEKAEADLKKAVDPE 99

RESULT 10
US-08-529-055-42
; Sequence 42, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-42
```



```
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-59

Query Match          90.3%; Score 437; DB 4; Length 188;
Best Local Similarity 89.9%; Pred. No. 1.7e-35;
Matches 89; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAAIEAEVQLKDAE 60
Db 1 LKEIDESDSEYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAAIEAEVQLKDAE 60

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99

RESULT 14
US-08-714-741-41
; Sequence 41, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: EXPRESSION AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1231 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid

US-08-714-741-41
Query Match          88.6%; Score 429; DB 4; Length 1231;
Best Local Similarity 90.7%; Pred. No. 1e-33;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 EINESDSEYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAAIEAEVQLKDAE 62
Db 494 EVLSESEDYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAAIEAEVQLKDAE 553

Qy 63 NNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 554 NNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 590

RESULT 15
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-286-981B-2

Query Match          88.4%; Score 428; DB 4; Length 141;
Best Local Similarity 90.7%; Pred. No. 9.2e-35;
Matches 88; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAAIEAEVQLKDAE 60
Db 45 LKEIDESDSEYAKGFRAPLQSKLDKAKKALLKLEELSGKIEELDAAIEAEVQLKDAE 104

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
Db 105 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 141

Search completed: June 21, 2005, 10:25:21
Job time : 19.4867 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-15

Perfect score: 484

Sequence: 1 LEEINESDEDYAKGFRAP.....KKAELKAEADLKKAVIDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 484   | 100.0       | 99     | 15 | US-10-674-755-15  |
| 2          | 454   | 93.8        | 206    | 15 | US-10-299-636-69  |
| 3          | 451   | 93.2        | 99     | 15 | US-10-674-755-14  |
| 4          | 444.5 | 91.8        | 100    | 15 | US-10-674-755-10  |
| 5          | 440   | 90.9        | 170    | 15 | US-10-299-636-75  |
| 6          | 440   | 90.9        | 181    | 15 | US-10-299-636-57  |
| 7          | 440   | 90.9        | 643    | 15 | US-10-299-636-95  |
| 8          | 440   | 90.9        | 670    | 9  | US-09-748-875-63  |
| 9          | 440   | 90.9        | 670    | 10 | US-09-298-523B-63 |
| 10         | 440   | 90.9        | 690    | 9  | US-09-748-875-61  |
| 11         | 440   | 90.9        | 690    | 10 | US-09-298-523B-61 |

|    |       |      |     |    |                      |                   |
|----|-------|------|-----|----|----------------------|-------------------|
| 12 | 440   | 90.9 | 691 | 9  | US-09-748-875-1      | Sequence 1, Appli |
| 13 | 440   | 90.9 | 691 | 10 | US-09-298-523B-1     | Sequence 1, Appli |
| 14 | 440   | 90.9 | 701 | 9  | US-09-748-875-62     | Sequence 62, Appl |
| 15 | 440   | 90.9 | 701 | 10 | US-09-298-523B-62    | Sequence 62, Appl |
| 16 | 440   | 90.9 | 707 | 9  | US-09-748-875-2      | Sequence 2, Appli |
| 17 | 440   | 90.9 | 707 | 10 | US-09-298-523B-2     | Sequence 2, Appli |
| 18 | 440   | 90.9 | 711 | 9  | US-09-748-875-3      | Sequence 3, Appli |
| 19 | 440   | 90.9 | 711 | 10 | US-09-298-523B-3     | Sequence 3, Appli |
| 20 | 440   | 90.9 | 739 | 17 | US-10-732-923-3294   | Sequence 3294, Ap |
| 21 | 440   | 90.9 | 929 | 9  | US-09-748-875-60     | Sequence 60, Appl |
| 22 | 440   | 90.9 | 929 | 10 | US-09-298-523B-60    | Sequence 60, Appl |
| 23 | 440   | 90.9 | 929 | 15 | US-10-299-636-94     | Sequence 94, Appl |
| 24 | 437   | 90.3 | 99  | 15 | US-10-674-755-16     | Sequence 16, Appl |
| 25 | 437   | 90.3 | 188 | 15 | US-10-299-636-74     | Sequence 74, Appl |
| 26 | 428   | 88.4 | 141 | 14 | US-10-254-995-2      | Sequence 2, Appli |
| 27 | 428   | 88.4 | 589 | 9  | US-09-748-875-14     | Sequence 14, Appl |
| 28 | 428   | 88.4 | 589 | 10 | US-09-298-523B-14    | Sequence 14, Appl |
| 29 | 428   | 88.4 | 589 | 15 | US-10-299-636-97     | Sequence 97, Appl |
| 30 | 426   | 88.0 | 204 | 15 | US-10-299-636-73     | Sequence 73, Appl |
| 31 | 401   | 82.9 | 198 | 15 | US-10-299-636-76     | Sequence 76, Appl |
| 32 | 401   | 82.9 | 354 | 15 | US-10-299-636-105    | Sequence 105, App |
| 33 | 401   | 82.9 | 588 | 15 | US-10-299-636-96     | Sequence 96, Appl |
| 34 | 401   | 82.9 | 619 | 10 | US-09-882-774-1      | Sequence 1, Appli |
| 35 | 401   | 82.9 | 619 | 15 | US-10-282-122A-73702 | Sequence 73702, A |
| 36 | 401   | 82.9 | 619 | 16 | US-10-414-532-72     | Sequence 72, Appl |
| 37 | 393   | 81.2 | 99  | 15 | US-10-674-755-11     | Sequence 11, Appl |
| 38 | 393   | 81.2 | 204 | 15 | US-10-299-636-66     | Sequence 66, Appl |
| 39 | 387.5 | 80.1 | 100 | 15 | US-10-674-755-12     | Sequence 12, Appl |
| 40 | 377   | 77.9 | 193 | 15 | US-10-299-636-64     | Sequence 64, Appl |
| 41 | 374   | 77.3 | 195 | 15 | US-10-299-636-86     | Sequence 86, Appl |
| 42 | 373   | 77.1 | 99  | 15 | US-10-674-755-17     | Sequence 17, Appl |
| 43 | 370   | 76.4 | 99  | 15 | US-10-674-755-13     | Sequence 13, Appl |
| 44 | 353.5 | 73.0 | 336 | 15 | US-10-299-636-103    | Sequence 103, App |
| 45 | 305.5 | 63.1 | 183 | 15 | US-10-299-636-65     | Sequence 65, Appl |

#### ALIGNMENTS

##### RESULT 1

US-10-674-755-15  
; Sequence 15, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-15

Query Match 100.0%; Score 484; DB 15; Length 99;  
Best Local Similarity 100.0%; Pred. No. 1.5e-33;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |   |   |    |
|----|---|---|----|
| Qy | 1 | LEEINESDEDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDAEIAELEVQLKDAE | 60 |
| Db | 1 | LEEINESDEDYAKGFRAPLQSKLDKAKKLLKLELSGKIELDAEIAELEVQLKDAE | 60 |

|    |    |  |    |
|----|----|--|----|
| Qy | 61 | GNNVNAVYFKGLEKTTAEKKAELKAEADLKKAVIDEPE | 99 |
| Db | 61 | GNNVNAVYFKGLEKTTAEKKAELKAEADLKKAVIDEPE | 99 |

##### RESULT 2





## US-09-298-523B-63

Query Match 90.9%; Score 440; DB 10; Length 670;  
Best Local Similarity 90.9%; Pred. No. 6.7e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDAAKAKLLKLELSGKIEELDAAIEALEVOLKDAE 60  
Db 498 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIEELDAAIEALEVOLKDAE 557  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 558 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 596

## RESULT 10

US-09-748-875-61  
; Sequence 61, Application US/09748875  
; Publication No. US20010016200A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/748,875  
; PRIOR FILING DATE: 2000-12-26  
; PRIOR APPLICATION NUMBER: 09/298,523  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-61

Query Match 90.9%; Score 440; DB 9; Length 690;  
Best Local Similarity 90.9%; Pred. No. 6.9e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDAAKAKLLKLELSGKIEELDAAIEALEVOLKDAE 60  
Db 529 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIEELDAAIEALEVOLKDAE 588  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 589 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 627

## RESULT 11

US-09-298-523B-61  
; Sequence 61, Application US/09298523B  
; Publication No. US20030059438A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/298,523B  
; CURRENT FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-298-523B-61

Query Match 90.9%; Score 440; DB 10; Length 690;  
Best Local Similarity 90.9%; Pred. No. 6.9e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDAAKAKLLKLELSGKIEELDAAIEALEVOLKDAE 60

Db 529 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIEELDAAIEALEVOLKDAE 588  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 589 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 627

## RESULT 12

US-09-748-875-1  
; Sequence 1, Application US/09748875  
; Publication No. US20010016200A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/748,875  
; PRIOR FILING DATE: 2000-12-26  
; PRIOR APPLICATION NUMBER: 09/298,523  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 691  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-1

Query Match 90.9%; Score 440; DB 9; Length 691;  
Best Local Similarity 90.9%; Pred. No. 6.9e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDAAKAKLLKLELSGKIEELDAAIEALEVOLKDAE 60  
Db 530 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIEELDAAIEALEVOLKDAE 589  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 590 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 628

## RESULT 13

US-09-298-523B-1  
; Sequence 1, Application US/09298523B  
; Publication No. US20030059438A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/298,523B  
; CURRENT FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 691  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-298-523B-1

Query Match 90.9%; Score 440; DB 10; Length 691;  
Best Local Similarity 90.9%; Pred. No. 6.9e-29;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 LEEINESDSEDYAKGFRAPLQSKLDAAKAKLLKLELSGKIEELDAAIEALEVOLKDAE 60  
Db 530 LKEIDESDSEDYKGLRAPLQSKLDTKAKLKLKLELSGKIEELDAAIEALEVOLKDAE 589  
Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 99  
Db 590 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAVDEPE 628



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-15  
Perfect score: 484  
Sequence: 1 LEEINSESDYAKGFRAP.....KKAELEKADLKKAVDEPE 99  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 401   | 82.9        | 619    | 2 A97887 | surface protein ps  |
| 2          | 401   | 82.9        | 619    | 2 A41971 | surface protein ps  |
| 3          | 121   | 25.0        | 744    | 2 P95013 | pneumococcal surfa  |
| 4          | 120.5 | 24.9        | 161    | 2 S48396 | tropomyosin TPM2 -  |
| 5          | 114.5 | 23.7        | 852    | 2 D72230 | conserved hypot het |
| 6          | 113.5 | 23.5        | 1006   | 2 C70445 | ATPase subunit of   |
| 7          | 110   | 22.7        | 233    | 2 S70531 | bbk2.11 protein pr  |
| 8          | 108.5 | 22.4        | 281    | 2 F75216 | hypothetical prote  |
| 9          | 102.5 | 21.2        | 650    | 2 A11333 | ABC transporter (A  |
| 10         | 102   | 21.1        | 1319   | 2 A28313 | glued protein - fr  |
| 11         | 101   | 20.9        | 369    | 2 AG1648 | hypothetical prote  |
| 12         | 100.5 | 20.8        | 229    | 2 S70532 | outer surface prot  |
| 13         | 100.5 | 20.8        | 2401   | 2 T28676 | rhostry protein -   |
| 14         | 100   | 20.7        | 880    | 2 F75103 | conserved hypot het |
| 15         | 98.5  | 20.4        | 399    | 2 E71169 | hypothetical prote  |
| 16         | 98.5  | 20.4        | 701    | 2 H98120 | choline binding pr  |
| 17         | 98.5  | 20.4        | 785    | 2 T01025 | hypothetical prote  |
| 18         | 98    | 20.2        | 1078   | 2 T18352 | protein P120 - Msc  |
| 19         | 97.5  | 20.1        | 1110   | 2 I51116 | NF-180 - sea lampr  |
| 20         | 97    | 20.0        | 1790   | 2 S67593 | transport protein   |
| 21         | 96    | 19.8        | 395    | 2 AC1754 | capsid protein [ba  |
| 22         | 96    | 19.8        | 650    | 2 AH1704 | ABC transporter (A  |
| 23         | 95.5  | 19.7        | 764    | 2 T05409 | hypothetical prote  |
| 24         | 94    | 19.4        | 635    | 2 A10625 | ABC transporter AT  |
| 25         | 94    | 19.4        | 2116   | 2 A26655 | myosin heavy chain  |
| 26         | 93.5  | 19.3        | 407    | 1 EDBRQ3 | immediate-early pr  |
| 27         | 93.5  | 19.3        | 935    | 2 T51930 | kinesin [imported]  |
| 28         | 93    | 19.2        | 2139   | 2 T18296 | myosin heavy chain  |
| 29         | 92.5  | 19.1        | 488    | 2 F97039 | hypothetical prote  |

|    |      |      |      |          |                     |
|----|------|------|------|----------|---------------------|
| 30 | 92.5 | 19.1 | 886  | 2 H69378 | conserved hypot het |
| 31 | 92   | 19.0 | 1169 | 2 A64505 | Flis homolog - Met  |
| 32 | 92   | 19.0 | 1173 | 2 T25539 | hypothetical prote  |
| 33 | 92   | 19.0 | 1177 | 2 B75150 | chromosome segrega  |
| 34 | 92   | 19.0 | 1269 | 2 F84730 | probable myosin he  |
| 35 | 92   | 19.0 | 1827 | 2 T16270 | hypothetical prote  |
| 36 | 91.5 | 18.9 | 403  | 2 T16171 | hypothetical prote  |
| 37 | 91.5 | 18.9 | 522  | 2 C98608 | hypothetical prote  |
| 38 | 91   | 18.8 | 646  | 2 AH1587 | bacteriophage prot  |
| 39 | 91   | 18.8 | 660  | 2 AC1652 | hypothetical prote  |
| 40 | 91   | 18.8 | 1392 | 2 A43336 | microtubule-vesicl  |
| 41 | 91   | 18.8 | 1427 | 2 S22695 | resin - human       |
| 42 | 90.5 | 18.7 | 419  | 2 G75062 | probable flagella-  |
| 43 | 90   | 18.6 | 433  | 2 A89951 | trigger factor lim  |
| 44 | 90   | 18.6 | 1133 | 2 T22976 | hypothetical prote  |
| 45 | 90   | 18.6 | 1138 | 2 T24635 | hypothetical prote  |

ALIGNMENTS

RESULT 1

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C;Accession: A97887

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: A97887

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-619 <XUR>

A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:AE007317; PIDN:AAK98925.1; PID:g

C;Genetics:

A;Gene: pspA

Query Match 82.9%; Score 401; DB 2; Length 619;

Best Local Similarity 84.8%; Pred. No. 7.9e-21;

Matches 84; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 LEEINSESDYAKGFRAPLQSKLDKAKKLLKLELSKIEELDAETAELESLVKLDAE 60

Db 223 LKEIDSESDYAKGFRAPLQSKLDKAKKLLKLELSKIEELDAETAELESLVKLDAE 282

QY 61 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99

Db 283 ENNVVEDYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 321

RESULT 2

A41971

surface protein pspA precursor - Streptococcus pneumoniae

N;Alternate names: pneumococcal surface protein A

C;Species: Streptococcus pneumoniae

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C;Accession: A41971; A60282; A33134

R;Yother, J.; Briles, D.E.

J. Bacteriol. 174, 601-609, 1992

A;Title: Structural properties and evolutionary relationships of PspA, a surface protein

A;Reference number: A41971; MUID:92105030; PMID:1729249

A;Accession: A41971

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-619 <YOT>

A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DRI0; GB:M74122; NID:q153840; PIDN:AAK2701

A;Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIIP:75636)

R;Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.





```
QY 62 NNNVEAYFKEGLEKTTAAKAELEKAEAD 90
Db 566 E-----TEAKRETLKEQRENDQLKSD 587

RESULT 6
C70445
ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: C70445
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70445
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1006 <AQF>
A:Cross-references: UNIPROT:O67588; GB:AE000750; NID:g2983999; PIDN:AA07550.1; PID:g298
A:Experimental source: strain VF5
C:Genetics:
A:Gene: clpB
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: hydrolase

Query Match 23.5%; Score 113.5; DB 2; Length 1006;
Best Local Similarity 36.3%; Pred. No. 1.2;
Matches 37; Conservative 22; Mismatches 24; Indels 19; Gaps 5;

QY 2 EINESDSE-DYAKGFRAPLOSKLDAKAKLK-LEELSGKTEELDAEIAEVLKDA 59
Db 557 EQIIIEANLKG DYKE-----AQLKIEAKLEKEKQELLGKGVGEAKIAELKKKIEB- 608

QY 60 EGNVNEAYFKEGLEKTTAAKAELE-----KAEADLKKAUDE 97
Db 609 -----LDEKIEAAERGDYKEAELEKIEAKLEKLEKLEQE 645

RESULT 7
C70531
bbk2.11 protein precursor - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S70531
R:Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norga
Mol. Microbiol. 18, 507-520, 1995
A:Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi oute
A:Reference number: S70531; MUID:96342380; PMID:8748034
A:Accession: S70531
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-233 <AKI>
A:Cross-references: UNIPROT:Q44739; EMBL:U30617; NID:g3309515; PIDN:AA046421.1; PID:g119
C:Superfamily: outer surface protein P ospF
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-233/Product: bbk2.11 protein #status predicted <MAT>

Query Match 22.7%; Score 110; DB 2; Length 233;
Best Local Similarity 31.7%; Pred. No. 0.5;
Matches 38; Conservative 20; Mismatches 36; Indels 26; Gaps 5;

QY 3 EINESDEDYAK-----EGFRAPLO-----SKLDAKAKLL--KLEELSGKTEELD 46'
Db 33 EQNLESSEQNKKTEQSIKKQVEGFLEILTDLKSLDKLDEKDTKEIBKQIQELKNKIEKLD 92

QY 47 AEIAELEV-----QLKDAEGNNVNEAYFKEGLEKTTAAKAELEKAEADLKKAUDE 97
Db 93 SKTSTIETSYSEYBEKINKIEKLKGKLEDKFIE-LEESLAKKKGERKKALQAKQKFE 151
```

## RESULT 8

```
F75216
hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F75216
R:Anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc
A:Reference number: A75001
A:Accession: F75216
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-281 <KAW>
A:Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49181
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2181
```

```
Query Match 22.4%; Score 108.5; DB 2; Length 281;
Best Local Similarity 34.8%; Pred. No. 0.77;
Matches 32; Conservative 20; Mismatches 23; Indels 17; Gaps 3;

QY 19 APLOSKLDAKAKLKLEELSGKTEELDAEIAEVLKDAEGN-----NNVEA 67
Db 153 AELESLEKAK-----EIEELKGKVEKLEQKKELEKAESEVKLMEYAKAKRAELEA 208

QY 68 YFKEGLEKTTAAKAELEKAEADLKKAUDEPE 99
Db 209 KUREVEEKVREE--ELERKVSLEERSLNEYE 238
```

## RESULT 9

```
A11333
ABC transporter (ATP-binding protein) homolog lmo2073 [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: A11333
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Faihi, H.,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
C:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11333
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-650 <GLA>
A:Cross-references: UNIPROT:Q8Y519; GB:NC_003210; PIDN:CAD00151.1; PID:g16411543; GSPDB:(
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2073
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
```

```
Query Match 21.2%; Score 102.5; DB 2; Length 650;
Best Local Similarity 32.7%; Pred. No. 4.6;
Matches 36; Conservative 21; Mismatches 32; Indels 21; Gaps 4;
```

```
QY 2 EINESDSEDYAKGFRPRA-----PLOSKLDAKAKLL-----KLEELSGKTEELDAEI 49
Db 541 KELARLDADRKKGQVEATASVRKLYNQEKEQKLLRQKRKRLEIEKSMETDEKI 600

QY 50 AELEVOLKDAEGNNVNEAYFKEGLEKTTAAKAELEKAEADLKKAUDEPE 99
Db 601 AELELQTNPE----VFQDHEKALEIT-----QELDAVTKADGKLMEEWE 641
```

## RESULT 10

```
A28313
glued protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
```

C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C;Accession: A28313  
R;Swaroop, A.; Swaroop, M.; Garen, A.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987  
A;Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued 9e  
A;Reference number: A28313; MUID:87317680; PMID:2819881  
A;Accession: A28313  
A;Molecule type: DNA; mRNA  
A;Residues: 1-1319 <SWA>  
A;Cross-references: UNIPROT:P13496  
A;Note: the authors' translation is inconsistent with the nucleotide sequence in the reg  
C;Genetics:  
A;Gene: FlyBase:Gl  
A;Cross-references: FlyBase:FBgn0001108  
A;Introns: 18/2; 479/3  
C;Keywords: cytoskeleton; glycoprotein  
F;397,590,771,888,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 21.1%; Score 102; DB 2; Length 1319;  
Best Local Similarity 31.7%; Pred. No. 10;  
Matches 32; Conservative 21; Mismatches 26; Indels 22; Gaps 4;  
Qy 1 LEEINESDSYAKGFRAPLQSKLDKAKKLLKLV-----BELSKIEELDAETAELEVL 56  
Db 429 LRDSAHDKHDIQK-----LSKELEMKRSEVTELTKEKLSAKIDELEAIVADLQEQV 482  
Qy 57 KDAEGNNVAYFKGEGLEKTTAEKKAELKAEADLKKAVDE 97  
Db 483 DAALG-----AEWVEQLAEKKWELE-----DKVKLLRE 511

RESULT 11  
AG1648  
hypothetical protein homolog lin1728 [imported] - Listeria innocua (strain Clip11262)  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AG1648  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vasquez-Boland, J.A.; Voss, H.; Wehlant,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AG1648  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-369 <GLA>  
A;Cross-references: UNIPROT:Q92B23; GB:AL592022; PIDN:CAC96959.1; PID:g16414215; GSPDB:G  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin1728

Query Match 20.9%; Score 101; DB 2; Length 369;  
Best Local Similarity 31.9%; Pred. No. 3.3;  
Matches 30; Conservative 18; Mismatches 36; Indels 10; Gaps 3;  
Qy 8 DSEDY-----AKGFRAPLQSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAEGN 62  
Db 198 DSKEYEYDVSVKSRFDA-LEAKDAANAANKAVDAGEKGLTKLDKAEASOLEEANKN 256  
Qy 63 NNVEAYFKGEGLEKTTAEKKAELKAEADLKKAVD 96  
Db 257 TMT----KEDIENAVTERALLDSAKVMIGDSFD 286

RESULT 12  
S70532  
outer surface protein F precursor - Lyme disease spirochete  
C;Species: Borrelia burgdorferi (Lyme disease spirochete)  
C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S70532

R;Akins, D.R.; Porcella, S.F.; Popova, T.G.; Shevchenko, D.; Baker, S.I.; Li, M.; Norgat  
Mol. Microbiol. 18, 507-520, 1995  
A;Title: Evidence for in vivo but not in vitro expression of a Borrelia burgdorferi outer  
A;Reference number: S70531; MUID:96342380; PMID:8748034  
A;Accession: S70532  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-229 <AKI>  
A;Cross-references: UNIPROT:Q44735; EMBL:U19754; NID:g3318660; PIDN:AAC26147.1; PID:g896  
C;Genetics:  
A;Gene: ospF  
C;Superfamily: outer surface protein F ospF  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-229/Product: outer surface protein F #status predicted <MAT>

Query Match 20.8%; Score 100.5; DB 2; Length 229;  
Best Local Similarity 33.9%; Pred. No. 2.3;  
Matches 38; Conservative 20; Mismatches 35; Indels 19; Gaps 6;  
Qy 3 EINESDSYAK---EGFRAPLQ-----SKLDKAKKLLKLEELSGKIEELDAE-----I 49  
Db 40 EQNVKTEQEIKKQVGFLEILTDLNKLDTKEIE-KRIQELKEKLEAKKTSLKTY 98  
Qy 50 AELEVLKLD-----AEGNNVAYFKGEGLEKTTAEKKAELKAEADLKKAVDE 97  
Db 99 SEYEELKQIKELKGLKADLEKLG-GLSDSLKKKKEERKKALEDKAKKPEE 149

RESULT 13  
T28676  
rhopty protein - Plasmodium yoelii (fragment)  
C;Species: Plasmodium yoelii  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T28676; A45521  
R;Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.  
Mol. Biochem. Parasitol. 76, 329-332, 1996  
A;Title: Comparison of two members of a multigene family coding for high-molecular mass  
A;Reference number: Z20507; MUID:97077455; PMID:8920022  
A;Accession: T28676  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-2401 <SIN>  
A;Cross-references: UNIPROT:Q26216; EMBL:U36927; NID:g1041784; PID:g1041785; PIDN:AAB412  
R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-246, 1990  
A;Title: Identification of the gene for a plasmodium yoelii rhopty protein. Multiple co  
A;Reference number: A45521; MUID:9110160; PMID:2270106  
A;Accession: A45521  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 2260-2401 <KEE>  
A;Cross-references: GB:M34281

Query Match 20.8%; Score 100.5; DB 2; Length 2401;  
Best Local Similarity 34.0%; Pred. No. 23;  
Matches 33; Conservative 14; Mismatches 45; Indels 5; Gaps 2;  
Qy 6 ESDESYAKGFRAPLQSKLDKAKKLLKLEELSGKIEE---LDAETAELEVLKDAEGN 62  
Db 165 DKSTYDEKKGFGFSSLEAKNWEKKLEITELUKKNEETVQLDIKIRELIKQIKLIEE 224  
Qy 63 NNVEAYFKGEGLEKTTAE--KKAELKAEADLKKAVDE 97  
Db 225 QKIVNDLKLELNKKIKSEITEKIEYIKKAVDLKKKEIEK 261

RESULT 14  
F75103  
conserved hypothetical protein PAB0812 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: F75103  
R;anonymous, Genoscope

Search completed: June 21, 2005, 10:12:01  
Job time : 10.9 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-15  
Perfect score: 484  
Sequence: 1 LEEINESDEYAKGFRAP.....KKAELEKADLKKAVIDEPE 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 484   | 100.0       | 395    | 2 Q9LAY2 | Q9lay2 streptococc |
| 2          | 484   | 100.0       | 408    | 2 Q9LAY0 | Q9lay0 streptococc |
| 3          | 465   | 96.1        | 224    | 2 Q8GNS8 | Q8gns8 streptococc |
| 4          | 462   | 95.5        | 249    | 2 Q9L575 | Q9l575 streptococc |
| 5          | 459   | 94.8        | 426    | 2 Q9LAY5 | Q9lay5 streptococc |
| 6          | 452   | 93.4        | 99     | 2 Q8KQK4 | Q8kqk4 streptococc |
| 7          | 440   | 90.9        | 739    | 2 Q9RQT4 | Q9rqt4 streptococc |
| 8          | 440   | 90.9        | 820    | 2 Q9RQT1 | Q9rqt1 streptococc |
| 9          | 440   | 90.9        | 929    | 2 Q9KK19 | Q9kk19 streptococc |
| 10         | 440   | 90.9        | 929    | 2 Q9ZAY5 | Q9zay5 streptococc |
| 11         | 427   | 88.2        | 437    | 2 Q9LAY4 | Q9lay4 streptococc |
| 12         | 401   | 82.9        | 619    | 2 Q54972 | Q54972 streptococc |
| 13         | 401   | 82.9        | 619    | 2 Q8DR10 | Q8dr10 streptococc |
| 14         | 400.5 | 82.7        | 869    | 2 Q9KK27 | Q9kk27 streptococc |
| 15         | 393   | 81.2        | 417    | 2 Q9LAY3 | Q9lay3 streptococc |
| 16         | 379   | 78.3        | 415    | 2 Q9LAY1 | Q9lay1 streptococc |
| 17         | 312.5 | 64.6        | 225    | 2 Q9L591 | Q9l591 streptococc |
| 18         | 311.5 | 64.4        | 394    | 2 Q9LAY6 | Q9lay6 streptococc |
| 19         | 311.5 | 64.4        | 395    | 2 Q9LAZ1 | Q9laz1 streptococc |
| 20         | 309.5 | 63.9        | 246    | 2 Q9L578 | Q9l578 streptococc |
| 21         | 306.5 | 63.3        | 222    | 2 Q9L577 | Q9l577 streptococc |
| 22         | 306.5 | 63.3        | 262    | 2 Q9L576 | Q9l576 streptococc |
| 23         | 306.5 | 63.3        | 415    | 2 Q9LAY7 | Q9lay7 streptococc |
| 24         | 303.5 | 62.7        | 393    | 2 Q9LAZ3 | Q9laz3 streptococc |
| 25         | 301.5 | 62.3        | 194    | 2 Q9LSB5 | Q9lsb5 streptococc |
| 26         | 301.5 | 62.3        | 218    | 2 Q6UBB2 | Q6ub2 streptococc  |
| 27         | 301.5 | 62.3        | 233    | 2 Q9L568 | Q9l568 streptococc |
| 28         | 301.5 | 62.3        | 236    | 2 Q9L569 | Q9l569 streptococc |
| 29         | 301.5 | 62.3        | 243    | 2 Q9L564 | Q9l564 streptococc |
| 30         | 301.5 | 62.3        | 243    | 2 Q9L567 | Q9l567 streptococc |
| 31         | 301.5 | 62.3        | 244    | 2 Q9L565 | Q9l565 streptococc |

|    |       |      |     |          |                    |
|----|-------|------|-----|----------|--------------------|
| 32 | 301.5 | 62.3 | 247 | 2 Q9L566 | Q9l566 streptococc |
| 33 | 301.5 | 62.3 | 249 | 2 Q9L570 | Q9l570 streptococc |
| 34 | 301.5 | 62.3 | 254 | 2 Q9L563 | Q9l563 streptococc |
| 35 | 301.5 | 62.3 | 401 | 2 Q9LAZ2 | Q9laz2 streptococc |
| 36 | 299.5 | 61.9 | 416 | 2 Q9LAY8 | Q9lay8 streptococc |
| 37 | 298.5 | 61.7 | 255 | 2 Q9L581 | Q9l581 streptococc |
| 38 | 298.5 | 61.7 | 255 | 2 Q9LSB6 | Q9lsb6 streptococc |
| 39 | 298.5 | 61.7 | 406 | 2 Q9LAZ0 | Q9laz0 streptococc |
| 40 | 287.5 | 59.4 | 340 | 2 Q8KQK5 | Q8kqk5 streptococc |
| 41 | 286.5 | 59.2 | 237 | 2 Q9L592 | Q9l592 streptococc |
| 42 | 286.5 | 59.2 | 395 | 2 Q9LAY9 | Q9lay9 streptococc |
| 43 | 277.5 | 57.3 | 207 | 2 Q8GNS9 | Q8gns9 streptococc |
| 44 | 178.5 | 36.9 | 653 | 2 Q34097 | Q34097 streptococc |
| 45 | 164   | 33.9 | 107 | 2 Q8KQK2 | Q8kqk2 streptococc |

ALIGNMENTS

RESULT 1

Q9LAY2

ID Q9LAY2 PRELIMINARY; PRT; 395 AA.

AC Q9LAY2;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE PspA (Fragment).

GN Name=pspA;

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=1313;

RN [1]

RP -SEQUENCE FROM N.A.

RC STRAIN=EF6796;

RX MEDLINE=20448953; PubMed=10992499;

RX DOI=10.1128/IAI.68.10.5889-5900.2000;

RA Hollingshead S.K.; Becker R.; Briles D.E.;

RT "Diversity of PspA: mosaic genes and evidence for past recombination in Streptococcus pneumoniae.";

RL Infect. Immun. 68:5889-5900(2000).

DR EMBL; AF071813; AAF2709.1; -

DR InterPro; IPR009053; Prefoldin.

DR InterPro; IPR011047; Qun\_abc\_DH\_like.

DR InterPro; IPR000533; Tropomyosin.

DR PRINTS; PR00194; TROPOMYOSIN.

FT NON\_TER 395

SQ SEQUENCE 395 AA; 42963 MW; 586EF956BCBCC1E CRC64;

Query Match 100.0%; Score 484; DB 2; Length 395;

Best Local Similarity 100.0%; Pred. No. 2.5e-23;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEEINESDEYAKGFRAPLOSKLDKAKKLLKLEELSGKTEELDAETAELEVLKDAE 60

Db 225 LEEINESDEYAKGFRAPLOSKLDKAKKLLKLEELSGKTEELDAETAELEVLKDAE 284

QY 61 GNNVVEAYFKGLEKTTAEKAELEKAEADLKKAVIDEPE 99

Db 285 GNNVVEAYFKGLEKTTAEKAELEKAEADLKKAVIDEPE 323

RESULT 2

Q9LAY0

ID Q9LAY0 PRELIMINARY; PRT; 408 AA.

AC Q9LAY0;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE PspA (Fragment).

GN Name=pspA;

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```

OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297BF CRC64;

Query Match 100.0%; Score 484; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDAKAKLLKLEELSGKIEELDAEIAELEVLKDAE 60
Db 228 LEEINESDSEYAKGFRAPLQSKLDAKAKLLKLEELSGKIEELDAEIAELEVLKDAE 287

Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 288 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 326

RESULT 3
Q8GNS8
ID Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2004 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -
DR HSP; P00192; IAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 1
FT NON_TER 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 96.1%; Score 465; DB 2; Length 224;
Best Local Similarity 94.9%; Pred. No. 2.4e-22;
Matches 94; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDAKAKLLKLEELSGKIEELDAEIAELEVLKDAE 60
Db 17 LKIDNESDSEYVKEGFRAPLQSELDTKKAKLLKLEELSGKIEELDAEIAELEVLKDAE 76

Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 77 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 115

RESULT 4

```

```

Q9L575
ID Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packiam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 95.5%; Score 462; DB 2; Length 249;
Best Local Similarity 94.9%; Pred. No. 4e-22;
Matches 94; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDAKAKLLKLEELSGKIEELDAEIAELEVLKDAE 60
Db 74 LKIDNESDSEYIKGFRAPLQSELDTKKAKLLKLEELSGKIEELDAEIAELEVLKDAE 133

Qy 61 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 134 GNNVVEAYFKGLEKTTAEKKAELKAEADLKKAVDEPE 172

RESULT 5
Q9L575
ID Q9L575 PRELIMINARY; PRT; 426 AA.
AC Q9L575
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=PspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -
DR HSP; P00192; IM6T.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR005533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426
FT NON_TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

```

|                       |        |                  |               |                                 |
|-----------------------|--------|------------------|---------------|---------------------------------|
| Query Match           | 94.8%  | Score 459;       | DB 2;         | Length 426;                     |
| Best Local Similarity | 93.9%; | Pred. No. 1e-21; | Mismatches 3; | Indels 0; Gaps 0;               |
| Matches               | 93;    | Conservative     | 3;            | Mismatches 3; Indels 0; Gaps 0; |

QY

1

LEEINESDSEDYAKGFRAPLQSXLDAKAKLLKEELSGKIEELDAAIEAEVQLKDAE

60

Db

215

LKDINESDSEDYVVEGLRAPLQSELDTKAKLLKEELSGKIEELDAAIEAEVQLKDAE

274

QY

61

GNNNVAYFKEGLEKTTAETKAEELEKAEADLKKAVDPE

99

Db

275

GNNNVAYFKEGLEKTTAETKAEELEKAEADLKKAVDPE

313

RESULT 6

Q8KQK4

ID Q8KQK4 PRELIMINARY; PRT; 99 AA.

AC Q8KQK4;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Pneumococcal surface protein A (Fragment).

GN Name=pSpA;

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=371/00;

RX MEDLINE=22170754; PubMed=12183557;

RX DOI=10.1128/IAI.70.9.5086-5090.2002;

RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C., Dias W.O., Leite L.C.C.;

RA "Analysis of serum cross-reactivity and cross-protection elicited by immunization with DNA vaccines against Streptococcus pneumoniae expressing pSpA fragments from different clades.";

RL Infect. Immun. 70:5086-5090(2002).

DR ENBL; AY082388; ALU92493.1; -.

FT NON\_TER

FT 1

FT 99

SQ SEQUENCE 99 AA; 11105 MW; 7AI3308CAI74A3A7 CRC64;

Query Match

93.4%; Score 452;

DB 2; Length 99;

Best Local Similarity

92.9%; Pred. No. 7.4e-22;

Matches

92; Conservative

4; Mismatches 3; Indels 0; Gaps 0;

QY

1

LEEINESDSEDYAKGFRAPLQSXLDAKAKLLKEELSGKIEELDAAIEAEVQLKDAE

60

Db

1

LKEIDESDSEDYVVEGLRAPLQSELDTKAKLLKEELSGKIEELDAAIEAEVQLKDAE

60

QY

61

GNNNVAYFKEGLEKTTAETKAEELEKAEADLKKAVDPE

99

Db

61

GNNNVAYFKEGLEKTTAETKAEELEKAEADLKKAVDPE

99

RESULT 7

Q9RQT4

ID Q9RQT4 PRELIMINARY; PRT; 739 AA.

AC Q9RQT4;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Hypothetical protein ppPc (Fragment).

GN Name=ppPc;

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=E134;

RX MEDLINE=20038319; PubMed=10569772;

```

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIEBLDAEIAELEVLKDAE 60
Db 530 LKEIDESDSEYKLEGLRAPLQSKLDTKAKKLSKLEELSDKIDELDAEIAKLEVLKDAE 589

Qy 61 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 9
Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=stf10;
RA MEDLINE=2188621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RX Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 90.9%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 3.4e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIEBLDAEIAELEVLKDAE 60
Db 530 LKEIDESDSEYKLEGLRAPLQSKLDTKAKKLSKLEELSDKIDELDAEIAKLEVLKDAE 589

Qy 61 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 10
Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RA MEDLINE=20038319; PubMed=10569772;

```

```

RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSSP; P06653; 1H8C.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 90.9%; Score 440; DB 2; Length 929;
Best Local Similarity 90.9%; Pred. No. 3.4e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIEBLDAEIAELEVLKDAE 60
Db 530 LKEIDESDSEYKLEGLRAPLQSKLDTKAKKLSKLEELSDKIDELDAEIAKLEVLKDAE 589

Qy 61 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 628

RESULT 11
Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW binding_1; 1.
DR NON_TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 88.2%; Score 427; DB 2; Length 437;
Best Local Similarity 88.9%; Pred. No. 1.1e-19;
Matches 88; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LEEINESDSEYAKGFRAPLQSKLDKAKKLLKLEELSGKIEBLDAEIAELEVLKDAE 60
Db 235 LKEIDESDSEYKLEGLRAPLQSKLDTKAKKLSKLEELSDKIDELDAEIAKLVVQLKDAE 294

Qy 61 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 99
Db 295 GNNNVEAYFKGLEKTTAAKKAELKAEADLKKAVDEPE 333

RESULT 12

```



```

Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name:pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_TaxID=N.A.
RP SEQUENCE FROM N.A.
RX MEDLINE=92105030; PubMed=1729249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis."
RL J. Bacteriol. 174:601-609(1992).
[2]
RP SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Signal.
FT SIGNAL.
FT CHAIN 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;
Query Match 82.9%; Score 401; DB 2; Length 619;
Best Local Similarity 84.8%; Pred. No. 7.1e-18;
Matches 84; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 LEEINSESDYAKGFRAPLQSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE 60
223 LKEIDSESDYAKGFRAPLQSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE 282

QY 61 GNNVVEAYFKEGKLETKTAETAEKAELEKAEADLKAVDEPE 99
283 ENNVEDYFKEGKLETKTAETAEKAELEKAEADLKAVNEPE 321

RESULT 13
Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name:pspA; OrderedLocusNames=spv0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]_TaxID=N.A.
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Albom W.E. Jr., Arnold J., Blaszcak L.C., BURGESS S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glaes J.S., Khoja H., Kraft A.-R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., Mcleaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,

```

```

RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Statrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6."
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Complete proteome.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;
Query Match 82.9%; Score 401; DB 2; Length 619;
Best Local Similarity 84.8%; Pred. No. 7.1e-18;
Matches 84; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 LEEINSESDYAKGFRAPLQSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE 60
223 LKEIDSESDYAKGFRAPLQSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE 282

QY 61 GNNVVEAYFKEGKLETKTAETAEKAELEKAEADLKAVDEPE 99
283 ENNVEDYFKEGKLETKTAETAEKAELEKAEADLKAVNEPE 321

RESULT 14
Q9KK27 PRELIMINARY; PRT; 869 AA.
AC Q9KK27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name-pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_TaxID=N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=G5;
RX MEDLINE=21188621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae."
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C-membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfam; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 869 AA; 98732 MW; AFF2B504347E0220 CRC64;
Query Match 82.7%; Score 400.5; DB 2; Length 869;
Best Local Similarity 84.8%; Pred. No. 1e-17;
Matches 84; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 1 LEEINSESDYAKGFRAPLQSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE 60
537 LKEIDSESDYAKGFRAPLQSKLDKAKKLLKLEELSGKIEELDAETAELEVLKDAE 595

QY 61 GNNVVEAYFKEGKLETKTAETAEKAELEKAEADLKAVDEPE 99
596 GNNVVEAYFKEGKLETKTAETAEKAELEKAEADLKAVNEPE 634

```

```
RESULT 15
Q9LAY3
ID Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSP; P00192; 256B.
FT NON TER 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match 81.2%; Score 393; DB 2; Length 417;
Best Local Similarity 82.8%; Pred. No. 1.6e-17;
Matches 82; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LEEINESDSDYAKGFRAPLOSKLDAKAKLLKLELSGKIELDIAEIAELEVLKDAE 60
Db 213 LKEIDESDSDYVKEGFRAPLQSELDKQAKLSKLELSDKIDELDAEIAKLEDLKAAE 272
Qy 61 GNNNVEAYFKGLEKTTAEKKAELKAEADLKKAVIDEPE 99
Db 273 ENNVVEDYFKGLEKTTAAKKAELKTEADLKKAVIDEPE 311
```

Search completed: June 21, 2005, 10:22:12  
Job time : 62.3194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-16

Perfect score: 485

Sequence: 1 LKIEDESDYLKGERAP.....KXAELEKARADLKXAVDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID      | Description        |
|------------|-------|-------------|--------|------------|--------------------|
| 1          | 477   | 98.4        | 170    | 7 ABW02614 | Abw02614 Rct135c p |
| 2          | 477   | 98.4        | 181    | 7 ABW02596 | Abw02596 0922134c  |
| 3          | 477   | 98.4        | 865    | 6 ABU08489 | Abu08489 S. pneumo |
| 4          | 477   | 98.4        | 929    | 2 AAW14593 | AAW14593 Streptoco |
| 5          | 477   | 98.4        | 929    | 2 AAY43384 | Aay43384 S. pneumo |
| 6          | 477   | 98.4        | 8991   | 6 ABU08487 | Abu08487 S. pneumo |
| 7          | 474   | 97.7        | 188    | 2 AAW14580 | AAW14580 Streptoco |
| 8          | 474   | 97.7        | 188    | 7 ABW02613 | Abw02613 Rct129c p |
| 9          | 465   | 95.9        | 588    | 6 ABU08491 | Abu08491 Coiled co |
| 10         | 465   | 95.9        | 589    | 2 AAY43392 | Aay43392 PspC alph |
| 11         | 463   | 95.5        | 204    | 2 AAW14578 | AAW14578 Streptoco |
| 12         | 463   | 95.5        | 204    | 7 ABW02612 | Abw02612 Rct123c p |
| 13         | 462.5 | 95.4        | 180    | 2 AAW14562 | AAW14562 Streptoco |
| 14         | 459.5 | 94.7        | 187    | 2 AAW14579 | AAW14579 Streptoco |
| 15         | 442   | 91.1        | 1231   | 6 ABU08490 | Abu08490 Fragment  |
| 16         | 440   | 90.7        | 206    | 2 AAW14574 | AAW14574 Streptoco |
| 17         | 440   | 90.7        | 206    | 7 ABW02608 | Abw02608 Db15c pne |
| 18         | 416   | 85.8        | 198    | 2 AAW14581 | AAW14581 Streptoco |
| 19         | 413   | 85.2        | 198    | 7 ABW02615 | Abw02615 Rx1c pneu |
| 20         | 413   | 85.2        | 315    | 2 AAY04375 | Aay04375 Streptoco |
| 21         | 413   | 85.2        | 619    | 2 AAR63437 | Aar63437 Pneumococ |
| 22         | 413   | 85.2        | 619    | 2 AAR87598 | Aar87598 Pneumococ |
| 23         | 413   | 85.2        | 619    | 2 AAR66911 | Aar66911 Pneumococ |
| 24         | 413   | 85.2        | 619    | 2 AAY41838 | Aay41838 Streptoco |
| 25         | 413   | 85.2        | 619    | 5 AAE18782 | AAe18782 S. pneumo |

|    |       |      |     |            |                     |
|----|-------|------|-----|------------|---------------------|
| 26 | 413   | 85.2 | 619 | 6 ABU45778 | Abu45778 Protein e  |
| 27 | 413   | 85.2 | 619 | 8 ADOS2126 | Ados2126 Streptoco  |
| 28 | 413   | 85.2 | 648 | 2 AAW70336 | Aaw70336 Pneumococ  |
| 29 | 413   | 85.2 | 648 | 2 AAW62274 | Aaw62274 Streptoco  |
| 30 | 413   | 85.2 | 648 | 2 AAY41837 | Aay41837 Streptoco  |
| 31 | 413   | 85.2 | 648 | 2 AAW87879 | Aaw87879 A. pneumoc |
| 32 | 413   | 85.2 | 653 | 2 AAW92456 | Aaw92456 S. pneumo  |
| 33 | 413   | 85.2 | 684 | 2 AAR73912 | Aar73912 Streptoco  |
| 34 | 411   | 84.7 | 204 | 2 AAW14571 | AAW14571 Streptoco  |
| 35 | 411   | 84.7 | 204 | 7 ABW02605 | Abw02605 Ef1019c p  |
| 36 | 393   | 81.0 | 653 | 2 AAR27150 | Aar27150 PspA frag  |
| 37 | 390.5 | 80.5 | 289 | 2 AAW62276 | Aaw62276 Streptoco  |
| 38 | 390.5 | 80.5 | 289 | 2 AAY41840 | Aay41840 Streptoco  |
| 39 | 390.5 | 80.5 | 289 | 2 AAW87910 | Aaw87910 Protein s  |
| 40 | 390.5 | 80.5 | 289 | 2 AAW92458 | Aaw92458 S. pneumo  |
| 41 | 386   | 79.6 | 195 | 2 AAW14591 | AAW14591 Streptoco  |
| 42 | 386   | 79.6 | 195 | 7 ABW02625 | Abw02625 Wu2c pneu  |
| 43 | 369   | 76.1 | 623 | 6 ABU08494 | Abu08494 Fragment   |
| 44 | 365   | 75.3 | 605 | 6 ABU08493 | Abu08493 Fragment   |
| 45 | 355   | 73.2 | 190 | 2 AAW14569 | Aaw14569 Streptoco  |

#### ALIGNMENTS

RESULT 1  
ABW02614  
ID ABW02614 standard; protein; 170 AA.

XX AC ABW02614;

XX DT 12-FEB-2004 (first entry)

XX DE Rct135c pneumococcal surface protein A (PspA) central region.

XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
XX KW immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX PN US6592876-B1.

XX PD 15-JUL-2003.

XX PF 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX WPI; 2003-862841/80.

XX Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

XX Example 6; SEQ ID NO 60; 121pp; English.

XX The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspAs) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies) or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Rct135c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention

XX  
 SQ Sequence 170 AA;

Query Match 98.4%; Score 477; DB 7; Length 170;  
 Best Local Similarity 99.0%; Pred. No. 2.5e-36;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDESDSYDKGGRAPLOSGLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60  
 DB 1 LKEIDESDSYDKGGRAPLOSGLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60

QY 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99  
 DB 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99

RESULT 2  
 ABW02596  
 ID ABW02596 standard; protein; 181 AA.

XX  
 AC ABW02596;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE 0922134c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX  
 DR WPI; 2003-862841/80.

Immunological composition for obtaining expression products used for  
 detecting the presence of Streptococcus pneumoniae or its strain,  
 comprises at least two different full length isolated gene encoding  
 pneumococcal surface protein A.

XX  
 PS Example 6; SEQ ID NO 42; 121pp; English.

XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (Pspas) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies), or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is 0922134c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention

XX  
 SQ Sequence 181 AA;

Query Match 98.4%; Score 477; DB 7; Length 181;  
 Best Local Similarity 99.0%; Pred. No. 2.7e-36;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDESDSYDKGGRAPLOSGLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60  
 DB 1 LKEIDESDSYDKGGRAPLOSGLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60

QY 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99  
 DB 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99

RESULT 3  
 ABU08489  
 ID ABU08489 standard; protein; 865 AA.

XX  
 AC ABU08489;  
 XX  
 DT 24-JUN-2003 (first entry)  
 XX  
 DE S. pneumoniae pneumococcal surface protein C (PspC) protein.  
 XX  
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..37  
 FT /label= Signal\_peptide  
 FT Protein 38..865  
 FT /label= Mature\_PspC\_protein  
 XX  
 PN US6500613-B1.  
 XX  
 PD 31-DEC-2002.  
 XX  
 PF 16-SEP-1996; 96US-00714741.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UVAL-) UNIV ALABAMA.  
 XX  
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 2003-361534/34.  
 DR N-PSDB; ABX95377.

Isolated PspC amino acid sequence used as polymerase chain reaction or  
 hybridization probe, comprises pneumococcal surface protein having alpha-  
 helical, proline rich and repeat regions.

XX  
 PS Claim 3; Fig 21; 186pp; English.

XX  
 CC The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents S. pneumoniae  
 CC PspC protein

XX  
 SQ Sequence 865 AA;

Query Match 98.4%; Score 477; DB 6; Length 865;  
 Best Local Similarity 99.0%; Pred. No. 1.8e-35;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 466 LKEIDSDSDYLKEGLRPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 525
QY 61 GNNVVEAYFKGLEKTTAAKAELEKAEADLKAVDEPE 99
DB 526 GNNVVEAYFKGLEKTTAAKAELEKAEADLKAVDEPE 564

RESULT 4
AAW14593
ID AAW14593 standard; protein; 929 AA.
XX
AC AAW14593;
XX
DT 17-OCT-2003 (revised)
DT 27-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspC.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain EF6796.
XX
FH Key
FT Peptide 1..37 Location/Qualifiers
FT Protein 38..929 /label= Sig_peptide
FT Domain 38..637 /label= Mat_protein
FT Region 41..242 /label= Alpha-helix
FT Region 41..242 /label= Repeat_1
FT Region 69..637 /notes= "alpha-helical repeat region"
FT Region 69..637 /label= Coiled-coil
FT Region 332..492 /notes= "breaks in 7-residue periodicity of the coiled-coil occur at amino acids 136-141, 261-304 and 383-387"
FT Region 332..492 /label= Repeat_2
FT Domain 627..689 /notes= "alpha-helical repeat region"
FT Domain 913..929 /label= Proline-rich
FT Domain 913..929 /label= C-terminal
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Vother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 1997-202002/18.
DR N-PSDB; AAT61728.
XX
PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
PS Disclosure; Fig 13; 296pp; English.
XX
CC This sequence comprises the pneumococcal protein surface C (pspC) of
CC Streptococcus pneumoniae strain EF6796. The sequence was deduced from the
CC pspC gene (AAT61728). Like PspA, PspC has an alpha-helical coiled-coil
CC region, proline-rich central region, repeat regions, with a choline
CC binding motif, and a C-terminal 17-amino acid tail. The 2 polypeptides
```

```
CC share 3 regions of high sequence identity. One is a protection-eliciting
CC region present within the alpha-helical domain. The others are the
CC proline-rich domain and a repeat domain shared with other choline-binding
CC proteins and thought to play a role in cell surface association. PspC and
CC PspA polypeptides, and their fragments, can be used in vaccines to
CC protect against S. pneumoniae infection and hence for the prevention of
CC diseases such as otitis media, meningitis, bacteraemia and pneumonia.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 929 AA;
XX
Query Match 98.4%; Score 477; DB 2; Length 929;
Best Local Similarity 99.0%; Pred. No. 1.9e-35;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 530 LKEIDSDSDYLKEGLRPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 589
QY 61 GNNVVEAYFKGLEKTTAAKAELEKAEADLKAVDEPE 99
DB 590 GNNVVEAYFKGLEKTTAAKAELEKAEADLKAVDEPE 628

RESULT 5
AAW43384
ID AAW43384 standard; protein; 929 AA.
XX
AC AAW43384;
XX
DT 27-JAN-2000 (first entry)
XX
DE S. pneumoniae PspC protein sequence.
XX
KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
KW epitopic region; immunogenic composition; vaccine composition; therapy;
KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;
KW anti-PspA antibody.
XX
OS Streptococcus pneumoniae.
XX
PN WO9593940-A1.
XX
PD 28-OCT-1999.
XX
PF 23-APR-1999; 99WO-US008895.
XX
PR 23-APR-1998; 98US-0082728P.
XX
PA (UYAL-) UNIV ALABAMA.
XX
PI Briles DE, Hollingshead SK, Brooks-Walter A;
XX
DR WPI; 1999-620581/53.
DR N-PSDB; AAZ31956.
XX
PT New epitope polypeptides of Pneumococcal surface protein C, used to
PT develop products for immunological, immunogenic or vaccine compositions,
PT particularly against Streptococcus pneumoniae infections.
XX
PS Example; Fig 11; 109pp; English.
XX
CC This sequence is the Streptococcus pneumoniae pneumococcal surface
CC protein C. The invention relates to an isolated and/or purified
CC polypeptide comprising at least one epitope or epitopic region of
CC Pneumococcal surface protein C (PspC). The polypeptides or vectors
CC containing sequence encoding them can be used as immunogenic,
CC immunological or vaccine compositions. The compositions can be used for
CC eliciting an immunological response against Streptococcus pneumoniae
CC (SP), which can cause otitis media, meningitis, bacteraemia and
CC pneumonia. They can be used for eliciting an anti-PspA antibody. The
CC nucleic acid molecules can also be used for detecting pspC, pspA or SP
XX
```

SQ Sequence 929 AA;

Query Match 98.4%; Score 477; DB 2; Length 929;  
 Best Local Similarity 99.0%; Pred. No. 1.9e-35;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60  
 DB LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 3726

QY 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
 DB 3727 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 3765

RESULT 7

AAW14580  
 ID AAW14580 standard; protein; 188 AA.  
 XX  
 AC AAW14580;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Rct135.  
 XX  
 PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 1997-202002/18.  
 XX  
 PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 PS  
 PS Example 6; Fig 13; 296pp; English.  
 XX  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct135.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 188 AA;

Query Match 97.7%; Score 474; DB 2; Length 188;  
 Best Local Similarity 98.0%; Pred. No. 5.4e-36;  
 Matches 97; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60  
 DB LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60

QY 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
 DB 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99

SQ Sequence 8991 AA;

Query Match 98.4%; Score 477; DB 6; Length 8991;  
 Best Local Similarity 99.0%; Pred. No. 2.9e-34;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ABU08487

ID ABU08487 standard; protein; 8991 AA.  
 XX  
 AC ABU08487;  
 XX  
 DT 24-JUN-2003 (first entry)  
 XX  
 DE S. pneumoniae pneumococcal surface protein A (PspA) protein.  
 XX  
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PH Key Location/Qualifiers  
 FT Misc-difference 1. .8991  
 FT /note= "All Xaa residues within this sequence are  
 FT unknown"

US6500613-B1.  
 31-DEC-2002.  
 16-SEP-1996; 96US-00714741.  
 15-SEP-1995; 95US-00529055.  
 (UVAL-) UNIV ALABAMA.  
 Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 Hollingshead S, Tart R, Brooks-Walter A;  
 WPI; 2003-361534/34.  
 Isolated PspC amino acid sequence used as polymerase chain reaction or  
 hybridization probe, comprises pneumococcal surface protein having alpha-  
 helical, proline rich and repeat regions.  
 Disclosure; Col 145-188; 186pp; English.  
 The present invention relates to the isolation of Streptococcus  
 pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 like protein having alpha-helical, proline rich and repeat regions. The  
 PspC and PspA proteins may be used in a vaccine to protect against  
 pneumococcal infections. The polynucleotide sequences encoding PspC and  
 PspA may be used for the expression of the proteins, and as PCR primers  
 or hybridisation probes. The present sequence represents S. pneumoniae  
 PspA protein

```

RESULT 8
ABW02613
ID AEW02613 standard; protein; 188 AA.
XX
XX AC ABW02613;
XX
XX DT 12-FEB-2004 (first entry)
XX
XX DE Rct129c pneumococcal surface protein A (PspA) central region.
XX
XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
XX OS Unidentified.
XX
XX PN US6592876-B1.
XX
XX PD 15-JUL-2003.
XX
XX PF 15-SEP-1995; 95US-00529055.
XX
XX PR 20-APR-1993; 93US-00048896.
XX
XX PR 06-JUN-1995; 95US-00465746.
XX
XX PA (UABR-) UAB RES FOUND.
XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX PT Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX pneumococcal surface protein A.
XX
XX PS Example 6; SEQ ID NO 59; 121pp; English.
XX
XX CC The present invention relates to an immunological composition comprising
XX at least 2 different full length isolated genes encoding pneumococcal
XX surface protein A (PspAs) from different groups based on restriction
XX fragment polymorphism analysis. The invention is useful for obtaining
XX expression products by recombinant techniques to detect, determine,
XX isolate or diagnose the presence of Streptococcus pneumoniae or its
XX strain. The expression product is useful for preparing antigenic,
XX immunological or vaccine compositions, for eliciting antibodies, an
XX immunological response (other than or additional to antibodies) or a
XX protective response (including antibody or other immunological response
XX by administering compositions to a host). The invention is also useful as
XX vaccines and in gene therapy. The present sequence is Rct129c
XX pneumococcal surface protein A (PspA) central region. This sequence is
XX used in the exemplification of the invention
XX
XX SQ Sequence 188 AA;

Query Match 97.7%; Score 474; DB 7; Length 188;
Best Local Similarity 98.0%; Pred. No. 5.4e-36;
Matches 97; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 60
DB 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 60
QY 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKAVDEPE 99
DB 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKAVDEPD 99

RESULT 9
ABU08491
ID ABU08491 standard; protein; 588 AA.
XX
XX AC ABU08491;
XX
XX DT 27-JAN-2000 (first entry)
XX
XX DE PspC alpha-helix coiled-coil region.
XX
XX KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
XX epitopic region; immunogenic composition; vaccine composition; therapy;
XX meningitis; immunological response; otitis media; bacteraemia; pneumonia;
XX anti-PspA antibody.
XX
XX OS Streptococcus pneumoniae.

```

---

```

XX
XX DT 24-JUN-2003 (first entry)
XX
XX DE Coiled coil motif of alpha-helix of S. pneumoniae PspC protein.
XX
XX KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;
XX alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;
XX antibacterial.
XX
XX OS Streptococcus pneumoniae.
XX
XX PN US6500613-B1.
XX
XX PD 31-DEC-2002.
XX
XX PF 16-SEP-1996; 96US-00714741.
XX
XX PR 15-SEP-1995; 95US-00529055.
XX
XX PA (UYAL-) UNIV ALABAMA.
XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 2003-361534/34.
XX
XX PT Isolated PspC amino acid sequence used as polymerase chain reaction or
XX hybridization probe, comprises pneumococcal surface protein having alpha-
XX helical, proline rich and repeat regions.
XX
XX PS Disclosure; Fig 23; 186pp; English.
XX
XX CC The present invention relates to the isolation of Streptococcus
XX pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX like protein having alpha-helical, proline rich and repeat regions. The
XX PspC and PspA proteins may be used in a vaccine to protect against
XX pneumococcal infections. The polynucleotide sequences encoding PspC and
XX PspA may be used for the expression of the proteins, and as PCR primers
XX or hybridization probes. The present sequence represents a coiled coil
XX motif of the alpha-helix of S. pneumoniae PspC protein
XX
XX SQ Sequence 588 AA;

Query Match 95.9%; Score 465; DB 6; Length 588;
Best Local Similarity 99.0%; Pred. No. 1.4e-34;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 60
DB 492 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAETAKLEVQLKDAE 551
QY 61 GNNVYAYFKEGLEKTTAEKKALEKAEADLKAVDE 97
DB 552 GNNVYAYFKEGLEKTTAEKKALEKAEADLKAVDE 588

RESULT 10
AAY43392
ID AAY43392 standard; protein; 589 AA.
XX
XX AC AAY43392;
XX
XX DT 27-JAN-2000 (first entry)
XX
XX DE PspC alpha-helix coiled-coil region.
XX
XX KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;
XX epitopic region; immunogenic composition; vaccine composition; therapy;
XX meningitis; immunological response; otitis media; bacteraemia; pneumonia;
XX anti-PspA antibody.
XX
XX OS Streptococcus pneumoniae.

```

```

XX PN W09953940-A1.
XX XX
XX PD 28-OCT-1999.
XX XX
XX PF 23-APR-1999; 99WO-US008895.
XX XX
XX PR 23-APR-1998; 98US-0082728P.
XX XX
XX PA (UYAL-) UNIV ALABAMA.
XX XX
XX PI Briles DE, Hollingshead SK, Brooks-Walter A;
XX XX
XX DR WPI; 1999-620581/53.
XX XX
XX PT New epitope polypeptides of Pneumococcal surface protein C, used to
XX XX
XX PT develop products for immunological, immunogenic or vaccine compositions,
XX XX
XX PT particularly against Streptococcus pneumoniae infections.
XX XX
XX PS Example 1; Fig 3; 109pp; English.
XX XX
XX CC This sequence is the coiled-coil region of the Streptococcus pneumoniae
XX XX
XX CC pneumococcal surface protein C. The invention relates to an isolated
XX XX
XX CC and/or purified polypeptide comprising at least one epitope or epitopic
XX XX
XX CC region of Pneumococcal surface protein C (PspC). The polypeptides or
XX XX
XX CC vectors containing sequence encoding them can be used as immunogenic,
XX XX
XX CC immunological or vaccine compositions. The compositions can be used for
XX XX
XX CC eliciting an immunological response against Streptococcus pneumoniae
XX XX
XX CC (SP), which can cause otitis media, meningitis, bacteraemia and
XX XX
XX CC pneumonia. They can be used for eliciting an anti-PspA antibody. The
XX XX
XX CC nucleic acid molecules can also be used for detecting pspC, pspA or SP
XX XX
XX SQ Sequence 589 AA;

Query Match 95.9%; Score 465; DB 2; Length 589;
Best Local Similarity 99.0%; Pred. No. 1.4e-34;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGGERAPLQSKLDTKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60
Db 493 LKEIDSDSDYLKEGGERAPLQSKLDTKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 552
Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
Db 553 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 589

RESULT 11
AAW14578
ID AAW14578 standard; protein; 204 AA.
XX XX
XX AC AAW14578;
XX XX
XX DT 17-OCT-2003 (revised)
XX DT 28-OCT-1997 (first entry)
XX XX
XX DE Streptococcus pneumoniae PspA central region.
XX XX
XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX KW bacteraemia; pneumonia.
XX XX
XX OS Streptococcus pneumoniae; strain Rct123.
XX XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 4 /note= "unidentified amino acid"
XX FT Misc-difference 8 /note= "unidentified amino acid"
XX FT
XX XX
XX PN W09709994-A1.
XX XX
XX PD 20-MAR-1997.
XX XX

```

```

PF 16-SEP-1996; 96WO-US014819.
XX XX
XX PR 15-SEP-1995; 95US-00529055.
XX XX
XX PA (UABR-) UAB RES FOUND.
XX XX
XX PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX PI Hollingshead S, Tart R, Brooks-Walter A;
XX XX
XX DR WPI; 1997-202002/18.
XX XX
XX PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX PT in vaccines for protecting animals against S.pneumoniae infection.
XX XX
XX PS Example 6; Fig 13; 296pp; English.
XX XX
XX CC This sequence shows the central portion, including the C-terminus of the
XX CC alpha-helix region and some of the proline-rich region, of pneumococcal
XX CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.
XX CC Comparison of the N-terminal and central regions (AAW14533-57 and
XX CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
XX CC be used to divide the strains into several families based on sequence
XX CC homologies. PspA polypeptides, or fragments of them, can be used in
XX CC vaccines to protect animals against S. pneumoniae infection and hence for
XX CC the prevention of diseases such as otitis media, meningitis, bacteraemia
XX CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
XX CC region and the immediate 5' tip of the coding sequence are likely to be
XX CC the critical sequences for predicting PspA cross-reactions and vaccine
XX CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX XX
XX SQ Sequence 204 AA;

Query Match 95.5%; Score 463; DB 2; Length 204;
Best Local Similarity 96.0%; Pred. No. 6.2e-35;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGGERAPLQSKLDTKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60
Db 1 IKEDSDXSSEDYLKEGLRAPLQSKLDTKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60
Qy 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 12
ABW02612
ID ABW02612 standard; protein; 204 AA.
XX XX
XX AC ABW02612;
XX XX
XX DT 12-FEB-2004 (first entry)
XX XX
XX DE Rct123c pneumococcal surface protein A (PspA) central region.
XX XX
XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX KW immunological; gene therapy; immunoestimulant.
XX XX
XX OS Unidentified.
XX XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 1. .204 /note= "Xaa = Unknown amino acid"
XX FT
XX XX
XX PN US6592876-B1.
XX XX
XX PD 15-JUL-2003.
XX XX
XX PF 15-SEP-1995; 95US-00529055.
XX XX
XX PR 20-APR-1993; 93US-00048896.
XX PR 06-JUN-1995; 95US-00465746.
XX XX

```



PA (UABR-) UAB RES FOUND.  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX WPI; 2003-862841/80.  
 XX Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX Example 6; SEQ ID NO 58; 121pp; English.  
 XX The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Rct123c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX Sequence 204 AA;  
 SQ  
 Query Match 95.5%; Score 463; DB 7; Length 204;  
 Best Local Similarity 96.0%; Pred. No. 6.2e-35;  
 Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 LKEIDSDSDYKKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60  
 Db 1 IKEXDSXSDYKKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60  
 QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKADEPE 99  
 Db 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKADEPE 99  
 RESULT 13  
 ID AAW14562  
 XX AAW14562 standard; protein; 180 AA.  
 AC AAW14562;  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 XX bacteriaemia; pneumonia.  
 KW Streptococcus pneumoniae; strain 0922134c.  
 XX WO9709994-A1.  
 XX 20-MAR-1997.  
 XX 16-SEP-1996; 96WO-US014819.  
 XX 15-SEP-1995; 95US-00529055.  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain 0922134c.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX Sequence 180 AA;  
 SQ  
 Query Match 95.4%; Score 462.5; DB 2; Length 180;  
 Best Local Similarity 98.0%; Pred. No. 5.9e-35;  
 Matches 97; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSDYKKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60  
 Db 1 LKEIDSDSDYKKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLKDAE 60  
 QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKADEPE 99  
 Db 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKADEPE 98  
 RESULT 14  
 ID AAW14579  
 XX AAW14579 standard; protein; 187 AA.  
 AC AAW14579;  
 XX 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX Streptococcus pneumoniae PspA central region.  
 DE PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteriaemia; pneumonia.  
 XX Streptococcus pneumoniae; strain Rct129.  
 OS WO9709994-A1.  
 XX 20-MAR-1997.  
 XX 16-SEP-1996; 96WO-US014819.  
 XX 15-SEP-1995; 95US-00529055.  
 XX (UABR-) UAB RES FOUND.  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX WPI; 1997-202002/18.  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX Example 6; Fig 13; 296pp; English.  
 XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal

CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct129.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 187 AA;

Query Match 94.7%; Score 459.5; DB 2; Length 187;  
 Best Local Similarity 97.0%; Pred. No. 1.2e-34;  
 Matches 96; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60  
 DB 1 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60  
 QY 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99  
 DB 61 GNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 98

## RESULT 15

ABU08490  
 ID ABU08490 standard; protein; 1231 AA.  
 XX  
 AC ABU08490;  
 XX  
 DT 24-JUN-2003 (first entry)  
 XX  
 DE Fragment of S. pneumoniae PspC protein.  
 XX  
 KW Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FN US6500613-B1.  
 XX  
 PD 31-DEC-2002.  
 XX  
 PF 16-SEP-1996; 96US-00714741.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UYAL-) UNIV ALABAMA.  
 XX  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 2003-361534/34.  
 XX

PT Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.  
 XX

PS Disclosure; Fig 22; 186pp; English.  
 XX

CC The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents a fragment of S.  
 CC pneumoniae PspC protein

XX  
 SQ Sequence 1231 AA;  
 Query Match 91.1%; Score 442; DB 6; Length 1231;  
 Best Local Similarity 65.8%; Pred. No. 4.7e-32;  
 Matches 98; Conservative 0; Mismatches 1; Indels 50; Gaps 1;  
 QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 54  
 DB 442 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSKDIDELDAEIAKLEVLQSESE 501  
 QY 55 -----QLKDAEGNNNVEAYFK 70  
 DB 502 DYAKGFRAPLQSKLDAKKAELKAEADLKKAVDEPE 99  
 QY 71 EGLEKTTAEKKAELKAEADLKKAVDEPE 99  
 DB 562 EGLEKTTAEKKAELKAEADLKKAVDEPE 590  
 Search completed: June 21, 2005, 10:10:15  
 Job time : 73.8459 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-16

Perfect score: 485

Sequence: 1 LKEIDSDSEDLKGERAP.....KKAELKAEADLKKAVIDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 485   | 100.0       | 99     | 4  | US-09-147-875A-16 |
| 2          | 477   | 98.4        | 170    | 4  | US-08-529-055-60  |
| 3          | 477   | 98.4        | 181    | 4  | US-08-529-055-42  |
| 4          | 477   | 98.4        | 864    | 4  | US-08-714-741-40  |
| 5          | 477   | 98.4        | 8991   | 4  | US-08-714-741-32  |
| 6          | 474   | 97.7        | 188    | 4  | US-08-529-055-59  |
| 7          | 465   | 95.9        | 141    | 4  | US-09-286-981B-2  |
| 8          | 465   | 95.9        | 588    | 4  | US-08-714-741-42  |
| 9          | 463   | 95.5        | 99     | 2  | US-08-710-749-15  |
| 10         | 463   | 95.5        | 204    | 4  | US-08-529-055-58  |
| 11         | 457   | 94.2        | 99     | 2  | US-08-710-749-17  |
| 12         | 456.5 | 94.1        | 100    | 4  | US-09-147-875A-10 |
| 13         | 448   | 92.4        | 99     | 4  | US-09-147-875A-14 |
| 14         | 442   | 91.1        | 1231   | 4  | US-08-714-741-41  |
| 15         | 440   | 90.7        | 206    | 4  | US-08-529-055-54  |
| 16         | 437   | 90.1        | 99     | 4  | US-09-147-875A-15 |
| 17         | 426   | 87.8        | 99     | 2  | US-08-710-749-13  |
| 18         | 423   | 87.2        | 99     | 2  | US-08-710-749-11  |
| 19         | 413   | 85.2        | 99     | 2  | US-08-710-749-11  |
| 20         | 413   | 85.2        | 198    | 4  | US-08-529-055-61  |
| 21         | 413   | 85.2        | 619    | 1  | US-08-465-746-2   |
| 22         | 413   | 85.2        | 619    | 1  | US-08-214-164-2   |
| 23         | 413   | 85.2        | 619    | 2  | US-08-467-852A-3  |
| 24         | 413   | 85.2        | 619    | 2  | US-08-246-636-2   |
| 25         | 413   | 85.2        | 619    | 2  | US-08-247-491A-3  |
| 26         | 413   | 85.2        | 619    | 2  | US-08-319-795-2   |
| 27         | 413   | 85.2        | 619    | 2  | US-08-468-985-2   |

```

28 413 85.2 619 3 US-08-312-949-2 Sequence 2, Appli
29 413 85.2 648 1 US-08-072-070-2 Sequence 2, Appli
30 413 85.2 648 1 US-08-469-434-2 Sequence 2, Appli
31 413 85.2 648 2 US-08-214-222-2 Sequence 2, Appli
32 413 85.2 648 2 US-08-467-852A-2 Sequence 2, Appli
33 413 85.2 648 2 US-08-468-718-2 Sequence 2, Appli
34 413 85.2 648 2 US-08-247-491A-2 Sequence 2, Appli
35 413 85.2 648 3 US-08-446-201-3 Sequence 3, Appli
36 413 85.2 695 1 US-08-127-499A-23 Sequence 23, Appli
37 413 85.2 695 1 US-08-482-847-23 Sequence 23, Appli
38 411 84.7 99 2 US-08-710-749-10 Sequence 10, Appli
39 411 84.7 99 4 US-09-147-875A-11 Sequence 11, Appli
40 411 84.7 204 4 US-08-529-055-51 Sequence 51, Appli
41 401 82.7 288 3 US-08-312-949-4 Sequence 4, Appli
42 401 82.7 288 3 US-08-446-201-4 Sequence 4, Appli
43 399.5 82.4 100 4 US-09-147-875A-12 Sequence 12, Appli
44 390.5 80.5 289 1 US-08-072-070-4 Sequence 12, Appli
45 390.5 80.5 289 1 US-08-469-434-4 Sequence 4, Appli

```

#### ALIGNMENTS

RESULT 1

US-09-147-875A-16  
; Sequence 16, Application US/09147875A

; Patent No. 6638516

; GENERAL INFORMATION:

; APPLICANT: BECKER et al.

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/09/147,875A

; CURRENT FILING DATE: 1999-05-24

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 16

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-147-875A-16

Query Match 100.0%; Score 485; DB 4; Length 99;

Best Local Similarity 100.0%; Pred. No. 5.1e-39; Indels 0; Gaps 0;

Matches 99; Conservative 0; Mismatches 0;

QY 1 LKEIDSDSEDLKGERAPLQSKLDTKKAELKAEADLKKAVIDEPE 99

Db 1 LKEIDSDSEDLKGERAPLQSKLDTKKAELKAEADLKKAVIDEPE 60

QY 61 GNNVVEAYFKEGLEKTTASKKAELKAEADLKKAVIDEPE 99

Db 61 GNNVVEAYFKEGLEKTTASKKAELKAEADLKKAVIDEPE 99

RESULT 2

US-08-529-055-60

; Sequence 60, Application US/08529055

; Patent No. 6592876

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Brooks-Walter, Alexis

; TITLE OF INVENTION: Pneumococcal Genes, Portions

; TITLE OF INVENTION: Thereof, Expression Products

; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue

; CITY: New York

```
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-60

Query Match 98.4%; Score 477; DB 4; Length 170;
Best Local Similarity 99.0%; Pred. No. 5.4e-38;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTTKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 1 LKEIDSDSDYLKEGERAPLQSKLDTTKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60

Qy 61 GNNVYAFYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVYAFYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 3
US-08-529-055-42
; Sequence 42, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-42

Query Match 98.4%; Score 477; DB 4; Length 181;
Best Local Similarity 99.0%; Pred. No. 5.9e-38;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTTKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
Db 1 LKEIDSDSDYLKEGERAPLQSKLDTTKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60

Qy 61 GNNVYAFYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 61 GNNVYAFYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 4
US-08-714-741-40
; Sequence 40, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USSS OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-40

Query Match      98.4%; Score 477; DB 4; Length 864;
Best Local Similarity 99.0%; Pred. No. 3.6e-37;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAETAKLEVQLKDAE 60
DB 465 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSDKIDELDAETAKLEVQLKDAE 524

QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
DB 525 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 563

RESULT 5
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: THEREFROM, AND USES OF SUCH GENES,
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-32

Query Match      98.4%; Score 477; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 5.6e-36;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAETAKLEVQLKDAE 60
DB 3667 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSDKIDELDAETAKLEVQLKDAE 3726
```

```
QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
DB 3727 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 3765

RESULT 6
US-08-529-055-59
; Sequence 59, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-59

Query Match      97.7%; Score 474; DB 4; Length 188;
Best Local Similarity 98.0%; Pred. No. 1.2e-37;
Matches 97; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAETAKLEVQLKDAE 60
DB 1 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLELSDKIDELDAETAKLEVQLKDAE 60

QY 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPE 99
DB 61 GNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVDEPD 99

RESULT 7
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
```

```
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-2

Query Match          95.9%; Score 465; DB 4; Length 141;
Best Local Similarity 99.0%; Pred. No. 6e-37;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYLKEGERAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
    |||||
Db 45 LKEIDESDSDYLKEGLRAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 104

Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAUDE 97
    |||||
Db 105 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAUDE 141

RESULT 8
US-08-714-741-42
; Sequence 42, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-42

Query Match          95.9%; Score 465; DB 4; Length 588;
Best Local Similarity 99.0%; Pred. No. 3.2e-36;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYLKEGERAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
    |||||
Db 492 LKEIDESDSDYLKEGLRAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 551

Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAUDE 97
    |||||
Db 552 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAUDE 588

RESULT 9
US-08-710-749-15
; Sequence 15, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-15

Query Match          95.5%; Score 463; DB 2; Length 99;
Best Local Similarity 96.0%; Pred. No. 6.2e-37;
Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKEIDESDSDYLKEGERAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
    |||||
Db 1 LKEIDESDSDYLKEGLRAPLQSKLDTKKAKLSKLELSKIDELDAEIAKLEVLKDAE 60

Qy 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAUDE 99
    |||||
Db 61 GNNVAYFKEGLEKTTAAKKAELKAEADLKKAUDE 99
```

```

/ STREET: 530 Fifth Avenue
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10036
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/710,749
/ FILING DATE: 20-SEP-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Frommer, William S.
/ REGISTRATION NUMBER: 25,506
/ REFERENCE/DOCKET NUMBER: 454312-2074
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 840-3333
/ TELEFAX: (212) 840-0712
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 99 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: n/a
/ TOPOLOGY: linear
/ MOLECULE TYPE: amino acid
/ US-08-710-749-17

Query Match          94.2%; Score 457; DB 2; Length 99;
Best Local Similarity 94.9%; Pred. No. 2.3e-36;
Matches 94; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKEIDESDSEYLYKGERAPLOSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 1 LKEIDESDSEYLYKGERAPLOSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60

QY 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 12
US-09-147-875A-10
/ Sequence 10, Application US/09147875A
/ Patent No. 6638516
/ GENERAL INFORMATION:
/ APPLICANT: BECKER et al.
/ TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
/ FILE REFERENCE: 454312-2471
/ CURRENT APPLICATION NUMBER: US/09/147,875A
/ CURRENT FILING DATE: 1999-05-24
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 10
/ LENGTH: 100
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
/ US-09-147-875A-10

Query Match          94.1%; Score 456.5; DB 4; Length 100;
Best Local Similarity 96.0%; Pred. No. 2.6e-36;
Matches 96; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 LKEIDESDSEYLYKGERAPLOSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDA 59
DB 1 LKEIDESDSEYLYKGERAPLOSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDA 60

QY 60 EGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
DB 61 EGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 100

```

```
RESULT 13
US-09-147-875A-14
; Sequence 14, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid
US-09-147-875A-14

Query Match          92.4%; Score 448; DB 4; Length 99;
Best Local Similarity 91.9%; Pred. No. 1.6e-35;
Matches 91; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYLKEGERAPLQSKLDTKKAALKLELSKIDELDAEIAKLEVLKDAE 60
Db 1 LKIDSDSDYLKEGERAPLQSKLDTKKAALKLELSKIDELDAEIAKLEVLKDAE 60

Qy 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99
Db 61 GNNVVEYFKEGLEKTTAEKKAELKAEADLKKAVIDEPE 99

RESULT 14
US-08-714-741-41
; Sequence 41, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1231 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-41

Query Match          91.1%; Score 442; DB 4; Length 1231;
Best Local Similarity 65.8%; Pred. No. 1.1e-33;
Matches 98; Conservative 0; Mismatches 1; Indels 50; Gaps 1;

Qy 1 LKIDSDSDYLKEGERAPLQSKLDTKKAALKLELSKIDELDAEIAKLEVLKDAE 54
Db 442 LKIDSDSDYLKEGERAPLQSKLDTKKAALKLELSKIDELDAEIAKLEVLKDAE 501
Qy 55 -----OLKDAEGNNNVEAYFK 70
Db 502 DYAKGFRAPLQSKLDAKKAALKLELSKIDELDAEIAKLEVLKDAE 561
Qy 71 EGLEKTTAEKKAELKAEADLKKAVIDEPE 99
Db 562 EGLEKTTAEKKAELKAEADLKKAVIDEPE 590

RESULT 15
US-08-529-055-54
; Sequence 54, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```



MOLECULE TYPE: peptide  
US-08-529-055-54

Query Match 90.7%; Score 440; DB 4; Length 206;  
Best Local Similarity 90.9%; Pred. No. 2.2e-34;  
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60  
Db 1 LKDIIDSDSDYAKEGLRAPLQSELDTKKAKLSKLEELSGKIEELDAETIXELEVQLKDAE 60  
QY 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99  
Db 61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDPE 99

Search completed: June 21, 2005, 10:25:21  
Job time : 18.4867 secs

**This Page Blank (uspto)**

```

Query Match      100.0%; Score 485; DB 15; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LKIDSDSDYLLKEGRAPLQSKLDTTKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60
      |||||
Db 1 LKIDSDSDYLLKEGRAPLQSKLDTTKAKLSKLELSKDIDELDAEIAKLEVLQKDAE 60
      |||||

Qy 61 GNNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVDEPE 99
      |||||
Db 61 GNNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVDEPE 99
      |||||

RESULT 2

```

US-10-299-636-75  
 ; Sequence 75, Application US/10299636  
 ; Publication No. US20040077847A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E  
 ; APPLICANT: McDaniel, Larry S  
 ; APPLICANT: Swiatlo, Edwin  
 ; APPLICANT: Yother, Janet  
 ; APPLICANT: Crain, Marilyn J  
 ; APPLICANT: Hollingshead, Susan  
 ; APPLICANT: Tart, Rebecca  
 ; APPLICANT: Brooks-Walter, Alexis  
 ; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
 ; FILE REFERENCE: 57909/361  
 ; CURRENT APPLICATION NUMBER: US/10/299,636  
 ; CURRENT FILING DATE: 2002-11-19  
 ; PRIOR APPLICATION NUMBER: 08/714,741  
 ; PRIOR FILING DATE: 1996-09-16  
 ; PRIOR APPLICATION NUMBER: 08/529,055  
 ; PRIOR FILING DATE: 1995-09-15  
 ; NUMBER OF SEQ ID NOS: 111  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 75  
 ; LENGTH: 170  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-10-299-636-75

Query Match 98.4%; Score 477; DB 15; Length 170;  
 Best Local Similarity 99.0%; Pred. No. 1.3e-31;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSKIDELDAEIAKLEVLQKDAE 60  
 Db 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSKIDELDAEIAKLEVLQKDAE 60  
 Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKA VDEPE 99  
 Db 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKA VDEPE 99

RESULT 3  
 US-10-299-636-57  
 ; Sequence 57, Application US/10299636  
 ; Publication No. US20040077847A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E  
 ; APPLICANT: McDaniel, Larry S  
 ; APPLICANT: Swiatlo, Edwin  
 ; APPLICANT: Yother, Janet  
 ; APPLICANT: Crain, Marilyn J  
 ; APPLICANT: Hollingshead, Susan  
 ; APPLICANT: Tart, Rebecca  
 ; APPLICANT: Brooks-Walter, Alexis  
 ; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
 ; FILE REFERENCE: 57909/361  
 ; CURRENT APPLICATION NUMBER: US/10/299,636  
 ; CURRENT FILING DATE: 2002-11-19  
 ; PRIOR APPLICATION NUMBER: 08/714,741  
 ; PRIOR FILING DATE: 1996-09-16  
 ; PRIOR APPLICATION NUMBER: 08/529,055  
 ; PRIOR FILING DATE: 1995-09-15  
 ; NUMBER OF SEQ ID NOS: 111  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 57  
 ; LENGTH: 181  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-10-299-636-57

Query Match 98.4%; Score 477; DB 15; Length 181;  
 Best Local Similarity 99.0%; Pred. No. 1.4e-31;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSKIDELDAEIAKLEVLQKDAE 60  
 Db 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSKIDELDAEIAKLEVLQKDAE 60  
 Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKA VDEPE 99  
 Db 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKA VDEPE 99

RESULT 4  
 US-10-299-636-95  
 ; Sequence 95, Application US/10299636  
 ; Publication No. US20040077847A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E  
 ; APPLICANT: McDaniel, Larry S  
 ; APPLICANT: Swiatlo, Edwin  
 ; APPLICANT: Yother, Janet  
 ; APPLICANT: Crain, Marilyn J  
 ; APPLICANT: Hollingshead, Susan  
 ; APPLICANT: Tart, Rebecca  
 ; APPLICANT: Brooks-Walter, Alexis  
 ; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
 ; FILE REFERENCE: 57909/361  
 ; CURRENT APPLICATION NUMBER: US/10/299,636  
 ; CURRENT FILING DATE: 2002-11-19  
 ; PRIOR APPLICATION NUMBER: 08/714,741  
 ; PRIOR FILING DATE: 1996-09-16  
 ; PRIOR APPLICATION NUMBER: 08/529,055  
 ; PRIOR FILING DATE: 1995-09-15  
 ; NUMBER OF SEQ ID NOS: 111  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 95  
 ; LENGTH: 643  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-10-299-636-95

Query Match 98.4%; Score 477; DB 15; Length 643;  
 Best Local Similarity 99.0%; Pred. No. 5.7e-31;  
 Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSKIDELDAEIAKLEVLQKDAE 60  
 Db 245 LKEIDSDSDYLKEGERAPLQSKLDTKKA KLSKLELSKIDELDAEIAKLEVLQKDAE 304  
 Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKA VDEPE 99  
 Db 305 GNNVVEAYFKEGLEKTTAAEKKAELEKAEADLKKA VDEPE 343

RESULT 5  
 US-09-748-875-63  
 ; Sequence 63, Application US/09748875  
 ; Publication No. US20010016200A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BRILES et al.  
 ; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
 ; FILE REFERENCE: 454312-3140  
 ; CURRENT APPLICATION NUMBER: US/09/748,875  
 ; CURRENT FILING DATE: 2000-12-26  
 ; PRIOR APPLICATION NUMBER: 09/298,523  
 ; PRIOR FILING DATE: 1999-04-23  
 ; NUMBER OF SEQ ID NOS: 78  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 63  
 ; LENGTH: 670  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-09-748-875-63

```
Query Match      98.4%; Score 477; DB 9; Length 670;
Best Local Similarity 99.0%; Pred. No. 5.9e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
Db 498 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 557

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 558 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 596

RESULT 6
US-09-298-523B-63
; Sequence 63, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-63

Query Match      98.4%; Score 477; DB 10; Length 670;
Best Local Similarity 99.0%; Pred. No. 5.9e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
Db 498 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 557

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 558 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 596

RESULT 7
US-09-748-875-61
; Sequence 61, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-61

Query Match      98.4%; Score 477; DB 9; Length 690;
Best Local Similarity 99.0%; Pred. No. 6.1e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
Db 529 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 588

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 590 LKEIDSDSDYLKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 8
US-09-298-523B-61
; Sequence 61, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-61

Query Match      98.4%; Score 477; DB 10; Length 690;
Best Local Similarity 99.0%; Pred. No. 6.1e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
Db 529 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 588

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 590 LKEIDSDSDYLKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 9
US-09-748-875-51
; Sequence 1, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; CURRENT FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-51

Query Match      98.4%; Score 477; DB 9; Length 691;
Best Local Similarity 99.0%; Pred. No. 6.1e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 60
Db 530 LKEIDSDSDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLKDAE 589

QY 61 GNNVYAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 590 LKEIDSDSDYLKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 10
```



```
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-2

Query Match      98.4%; Score 477; DB 10; Length 707;
Best Local Similarity 99.0%; Pred. No. 6.3e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLYKEGERAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
Db 530 LKEIDSDSDYLYKEGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLQKDAE 589

QY 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 590 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 628

RESULT 15
US-09-748-875-3
; Sequence 3, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; TITLE OF INVENTION: AND STRAINS THEREOF AND USES THEREFOR
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298,523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-3

Query Match      98.4%; Score 477; DB 9; Length 711;
Best Local Similarity 99.0%; Pred. No. 6.3e-31;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLYKEGERAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
Db 539 LKEIDSDSDYLYKEGLRAPLQSKLDTKKAQKLSKLEELSDKIDELDAEIAKLEVLQKDAE 598

QY 61 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 599 GNNVAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 637
```

Search completed: June 21, 2005, 11:18:36  
Job time : 64.2388 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-16

Perfect score: 485

Sequence: 1 LKEIDSESDYLKEGERAP.....KKAELEKAEADLKKAVDEPE 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*\*

1: piri:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 413   | 85.2        | 619    | 2 A97887 | surface protein ps  |
| 2          | 413   | 85.2        | 619    | 2 A41971 | surface protein ps  |
| 3          | 124.5 | 25.7        | 744    | 2 P95013 | pneumococcal surfa  |
| 4          | 121.5 | 25.1        | 161    | 2 S48396 | tropomyosin TPM2 -  |
| 5          | 115.5 | 23.8        | 1818   | 1 S73852 | hypothetical prote  |
| 6          | 110.5 | 22.8        | 852    | 2 D72230 | conserved hypothet  |
| 7          | 105   | 21.6        | 1177   | 2 B75150 | chromosome segrega  |
| 8          | 104   | 21.4        | 743    | 2 D84854 | hypothetical prote  |
| 9          | 104   | 21.4        | 896    | 2 S43074 | epidermal growth f  |
| 10         | 104   | 21.4        | 1189   | 2 A64505 | Fill5 homolog - Met |
| 11         | 103.5 | 21.3        | 1006   | 2 C70445 | ATPase subunit of   |
| 12         | 103   | 21.2        | 281    | 2 F75216 | hypothetical prote  |
| 13         | 103   | 21.2        | 1319   | 2 A28313 | glued protein - fr  |
| 14         | 103   | 21.2        | 2116   | 2 A26655 | myosin heavy chain  |
| 15         | 102.5 | 21.1        | 650    | 2 A11333 | ABC transporter (A  |
| 16         | 102   | 21.0        | 395    | 2 AC1754 | capsid protein (ba  |
| 17         | 102   | 21.0        | 764    | 2 T05409 | hypothetical prote  |
| 18         | 101.5 | 20.9        | 229    | 2 S70532 | outer surface prot  |
| 19         | 101.5 | 20.9        | 1093   | 2 S48460 | probable membrane   |
| 20         | 101.5 | 20.9        | 1827   | 2 T16270 | hypothetical prote  |
| 21         | 101   | 20.8        | 1179   | 2 F71190 | probable chromosom  |
| 22         | 101   | 20.8        | 1132   | 2 T30845 | probable DNA repla  |
| 23         | 101   | 20.8        | 2139   | 2 T18296 | myosin heavy chain  |
| 24         | 100.5 | 20.7        | 233    | 2 S70531 | bbk2.11 protein pr  |
| 25         | 100.5 | 20.7        | 629    | 2 F86351 | protein T26F17.2 [  |
| 26         | 100.5 | 20.7        | 886    | 2 H69378 | conserved hypothet  |
| 27         | 100   | 20.6        | 880    | 2 F75103 | conserved hypothet  |
| 28         | 99    | 20.4        | 387    | 2 S57834 | fcrA protein precu  |
| 29         | 99    | 20.4        | 388    | 2 A46173 | Mrp4 protein - Str  |

hypothetical prote  
M protein precursor  
protein F32E10.3 [NF-180 - sea lamp  
fcrA 15 protein -  
Fc gamma (IgG) rec  
hypothetical prote  
protein P120 - Myc  
myosin heavy chain  
hypothetical prote  
hypothetical prote  
fcrA protein precu  
hypothetical prote  
hypothetical prote  
probable zuotin [1  
bacteriophage prot

30 99 20.4 1979 2 C71622  
31 98.5 20.3 501 2 A44643  
32 98.5 20.3 990 2 H88733  
33 98.5 20.3 1110 2 I51116  
34 98 20.2 388 2 S52536  
35 98 20.2 405 2 A33939  
36 98 20.2 638 2 C64156  
37 98 20.2 1078 2 T18352  
38 98 20.2 1976 2 A59252  
39 97.5 20.1 1053 2 T51375  
40 97.5 20.1 3488 2 T34418  
41 97 20.0 415 2 S35760  
42 97 20.0 1133 2 T22976  
43 96.5 19.9 385 2 T20410  
44 96.5 19.9 445 2 T50972  
45 96.5 19.9 646 2 AH1587

ALIGNMENTS

RESULT 1  
A97887  
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; Lealanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl  
C:Genetics:  
A:Gene: pspA

Query Match 85.2%; Score 413; DB 2; Length 619;  
Best Local Similarity 87.9%; Pred. No. 1.1e-20;  
Matches 87; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDSESDYLKEGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 60  
DB 223 LKEIDSESDYLKEGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLQKDAE 282

QY 61 GNNVVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
DB 283 ENNVVDYFKEGLEKTTAEKAELEKAEADLKKAVNEPE 321

RESULT 2  
A41971  
surface protein pspA precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A33134  
R:Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:1729249  
A:Accession: A41971  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:q153840; PIDN:AAA2701  
A:Note: sequence extracted from NCBI backbone (NCIN:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

```
Infect. Immun. 59, 1285-1289, 1991
A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability
A;Reference number: A60282; MUID:91169598; PMID:2004810
A;Accession: A60282
A;Molecule type: protein
A;Residues: 32-76 <TAL>
A;Experimental source: strain JY2008
C;Genetics:
A;Gene: pspA
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-619/Product: surface protein pspA #status predicted <MAT>
F;411-430/Domain: cpl repeat homology <CP01>
F;431-450/Domain: cpl repeat homology <CP02>
F;451-470/Domain: cpl repeat homology <CP03>
F;471-490/Domain: cpl repeat homology <CP04>
F;491-510/Domain: cpl repeat homology <CP05>
F;511-530/Domain: cpl repeat homology <CP06>
F;531-550/Domain: cpl repeat homology <CP07>
F;551-570/Domain: cpl repeat homology <CP08>
F;571-591/Domain: cpl repeat homology <CP09>
F;592-611/Domain: cpl repeat homology <CP10>

Query Match      85.2%; Score 413; DB 2; Length 619;
Best Local Similarity 87.9%; Pred. No. 1.1e-20;
Matches 87; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LKETDESSEDYLKKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 LKEIDESSEDYAKGFRAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLELDQKAAE 282

Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELKAEADLKKAVDEPE 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 ENNVVEDYFKEGLEKTTAAEKKAELKAEADLKKAVNEPE 321

RESULT 3
F95013
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95013
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid-
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-744 <KUR>
A;Cross-references: UNIPROT:Q97T39; GB:AB005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0117

Query Match      25.7%; Score 124.5; DB 2; Length 744;
Best Local Similarity 33.3%; Pred. No. 0.28;
Matches 44; Conservative 14; Mismatches 33; Indels 41; Gaps 5;

Qy 2 KEIDE-----SPSEDYLKKEGERAPLQSKLDTKKAKLSK-----LEELSDKI-----D 43
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 314 KEISNLBILGGADPED-----DTAALQNKLAAKAEAKKQTFLEKLLSLDPEGKTQD 368

Qy 44 ELDAEIAKLEVLQKDAAGNNVVEAYFKE-----GLEKTTAEKKAEL 85
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 369 ELDKAEAEALDKADELQNKVADLEKEISNLEILGGADSEDDTAALQNKLATKKAEL 428

Qy 86 KAEADLKKAVDE 97
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 429 KTKQELDAALNE 440
```

## RESULT 4

```
S48396
tropomyosin TPM2 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein Yili138c
C;Species: Saccharomyces cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S48396; A56490
R;Churcher, C.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48310
A;Accession: S48396
A;Molecule type: DNA
A;Residues: 1-161 <CHU>
A;Cross-references: UNIPROT:P40414; GB:247047; EMBL:Z38059; NID:g603997; PID:g763208; MI
R;Drees, B.; Brown, C.; Barrell, B.G.; Bretecher, A.
J. Cell Biol. 128, 383-392, 1995
A;Title: Tropomyosin is essential in yeast, yet the TPM1 and TPM2 products perform disti
A;Reference number: A56490; MUID:95146545; PMID:7844152
A;Accession: A56490
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-161 <DRE>
A;Cross-references: GB:247047; GB:Z38059; NID:g603997; PID:g763208
C;Genetics:
A;Gene: SGD:TPM2
A;Cross-references: SGD:S0001400; MIPS:YIL138c
A;Map position: 9L
C;Superfamily: tropomyosin TPM1
C;Keywords: cytoskeleton

Query Match      25.1%; Score 121.5; DB 2; Length 161;
Best Local Similarity 33.7%; Pred. No. 0.097;
Matches 34; Conservative 21; Mismatches 21; Indels 25; Gaps 4;

Qy 1 LKEIDESSEDYLKKEGERAPLQSKLDTKKAKLSKLELSDKIDELDAEIAKLEVLQKDAE 58
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 29 LKELEQSNT-----IKLSAKNEQLDSEVEKLELSQSDTK 68

Qy 59 --REGNNVVEAYFKEGLEKTTAEKKAELKAEADLKKAVDE 97
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 69 QLAEDSNLNRSN--NENYTKNQDLEQLEEDSEAKLKEAMDK 108

RESULT 5
S73852
hypothetical protein MG218 homolog F10_orf1818 - Mycoplasma pneumoniae (strain ATCC 2934:
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S73852
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk1, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73852
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1818 <HIM>
A;Cross-references: UNIPROT:P75471; EMBL:AE000051; GB:U00089; NID:gl674211; PIDN:AA89617
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: Mycoplasma genitalium hypothetical protein MG218

Query Match      23.8%; Score 115.5; DB 1; Length 1818;
Best Local Similarity 28.7%; Pred. No. 2.7;
Matches 33; Conservative 22; Mismatches 31; Indels 29; Gaps 4;

Qy 14 KEGERAPLQSKLDTKKAKL-----SKLEELSDKIDELDAEIAKLEVLQK---DAEGNNN-- 64
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1364 KEGELOGIQLKLSLKKTKQTBQEFKSLYQQRKLDRTQRTTSLKHLRELKAQNEATAHKNRE 1423
```

```
QY 65 ---VEAYFKEGLEKTTAEK-----KAELEKAEADLKKAVDEPE 99
Db 1424 VLEIENYKKELQRLTTEKSEFDNNKQRLFEYFKIRNEIEKKAHIKTVLEETQ 1478

RESULT 6
D72230
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: D72230
R:Nelson, K.B.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72230
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852 <ARN>
A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3670
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1636
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 22.8%; Score 110.5; DB 2; Length 852;
Best Local Similarity 31.5%; Pred. No. 2.7;
Matches 28; Conservative 20; Mismatches 30; Indels 11; Gaps 2;

QY 6 ESDSDYLKEGRAPLQSKLDTKKAKLSKLE-----LSDKIDELDAETAKLEVLQKDAEG 61
Db 506 EKIEELHRLGYSDLEELKDEKRLKRIEERHSISQKITAADVQISQIENQLKEIKG 565

QY 62 NNNVEAYFKEGLEKTTAEKAELEKAEAD 90
Db 566 E-----TEAKRETLEQREEMDLKSD 587

RESULT 7
B75150
chromosome segregation protein (smc1) PAB2109 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B75150
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: B75150
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1177 <KAW>
A:Cross-references: UNIPROT:Q9V1R8; GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB4928
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2109
C:Superfamily: chromosome segregation protein SMC1

Query Match 21.6%; Score 105; DB 2; Length 1177;
Best Local Similarity 31.1%; Pred. No. 8.8;
Matches 33; Conservative 26; Mismatches 35; Indels 12; Gaps 4;

QY 2 KEIDESDYLKEGRAPLQSKLDTKKAKL-----SKLELSDKIDELDAETAKLEVLQK 57
Db 403 KSLYENADIKRLAEKRLSRITLTKALPGIREVEVKRLKLEKKAELSNVENKIS 462

QY 58 D-AEGNNNVEAYFKEGLEKTTAEK---AELEKAEADLKKAVDEPE 99
Db 463 SISQRRKVE-----EELEKKTSELOKVSSELSLESLRELKAEQAQSE 504
```

```
RESULT 8
D84854
hypothetical protein At2g42480 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84854
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-769, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84854
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-743 <STO>
A:Cross-references: UNIPROT:Q9SLB2; GB:AE002093; NID:g4567318; PIDN:AAD23729.1; GSPDB:GN
C:Genetics:
A:Gene: At2g42480
A:Map position: 2

Query Match 21.4%; Score 104; DB 2; Length 743;
Best Local Similarity 34.9%; Pred. No. 6.5;
Matches 30; Conservative 16; Mismatches 30; Indels 10; Gaps 3;

QY 6 ESDSDYLKEGRAPLQSKLDTKKAKLSKLE-ELSDKIDELDAETAKLEVLQKDAEGNNN 64
Db 260 EQDIEERLKNLEGMFEFDSKLSKLDLISLERKKSVDADGSRVQQLSERVKD----- 313

QY 65 VEAYFKEGLEKTTAEKAELEKAEAD 90
Db 314 IELILKSKLEEVSSSEKK---KKADAD 336

RESULT 9
S43074
epidermal growth factor receptor substrate - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43074; I38525
R:Bernard, O.A.; Mauchauffe, M.; Mecucci, C.; van den Berghe, H.; Berger, R.
Oncogene 9, 1039-1045, 1994
A:Title: A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AF-4, i
A:Reference number: S43074; MUID:94181254; PMID:8134107
A:Accession: S43074
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-896 <BER>
A:Cross-references: UNIPROT:P42566; EMBL:Z29064; NID:g470034; PIDN:CAA82305.1; PID:g47003
R:Wong, W.T.; Kraus, M.H.; Carlomagno, F.; Zelano, A.; Druck, T.; Croce, C.M.; Huebner, I
Oncogene 9, 1591-1597, 1994
A:Title: The human eps15 gene, encoding a tyrosine kinase substrate, is conserved in evo
A:Reference number: I38525; MUID:94239734; PMID:8183552
A:Accession: I38525
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-821, 'M', 823-896 <RES>
A:Cross-references: EMBL:U07707; NID:g466259; PIDN:AAA52101.1; PID:g466260
C:Genetics:
A:Gene: GDB:EPS15; AF-1P; MLT5
A:Cross-references: GDB:360337; OMIM:600051
A:Map position: lp32-1p32

Query Match 21.4%; Score 104; DB 2; Length 896;
Best Local Similarity 28.7%; Pred. No. 7.8;
Matches 29; Conservative 22; Mismatches 44; Indels 6; Gaps 2;

QY 3 EIDESDSELYLKE--GERAPLQSKLDTKKAKLSKLE-----EELSCKIDELDAETAKLEVLQ 56
Db 353 EODLXEKEDTIKQRTSEVQDLQDEVRENTNQLKQAQKQVQQLDELDELDEQKAQLEQL 412

QY 57 KDAEGNNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDE 97
```

Db 413 KEVRKKCAEAQLISSLKABLTQSQISTYEBELAKAREE 453

RESULT 10

A64505

P115 homolog - Methanococcus jannaschii

C;Species: Methanococcus jannaschii

C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 02-Jun-2000

C;Accession: A64505

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinscock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A;Reference number: A64300; MUID:96337999; PMID:8698087

A;Accession: A64505

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1169 <BUL>

A;Cross-references: GB:U67604; GB:L77117; NID:gl592224; PID:gl500543; TIGR:MJ1643

C;Genetics:

A;Map position: FOR1623481-1626990

C;Superfamily: chromosome segregation protein SMC1

Query Match 21.4%; Score 104; DB 2; Length 1169;

Best Local Similarity 27.1%; Pred. No. 10;

Matches 29; Conservative 27; Mismatches 33; Indels 18; Gaps 3;

Qy 1 LKSIDSDSDYLKGERAPLOSGLD-----TKAKLSKLELSKDIDELDAIAKLEVQ 55

Db 799 LKRWNETEGELKILEKAKLKNEDKGLTLVKEILPKIEELNKKVSELINKKVLKLN 858

Qy 56 LKDAEGNNVNEAFKGELEKTTA---EKKALEKAEADLKAVDEPE 99

Db 859 I-----SPYKESIEKNLSILEEKRYRYBELAKNLKELTEKKE 895

RESULT 11

C70445

ATPase subunit of ATP-dependent proteinase (EC 3.4.-.-) - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004

C;Accession: C70445

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: C70445

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1006 <AQF>

A;Cross-references: UNIPROT:O67588; GB:AE000750; NID:g2983999; PIDN:AAC07550.1; PID:g298

A;Experimental source: strain VF5

C;Genetics:

A;Gene: clpB

C;Superfamily: endopeptidase Clp ATP-binding chain

C;Keywords: hydrolase

Query Match 21.3%; Score 103.5; DB 2; Length 1006;

Best Local Similarity 32.7%; Pred. No. 9.5;

Matches 35; Conservative 22; Mismatches 27; Indels 23; Gaps 5;

Qy 1 LKSIDSDSDYLKGERAPLOSGLD-----AQLKIEKAKLEKQELGLGVGVGAIAELKK 604

Db 552 IKALEEQIIEANLKGDIYEK-----AQLKIEKAKLEKQELGLGVGVGAIAELKK 604

Qy 55 QLKDAEGNNVNEAFKGELEKTTAETKKALELE---KAEADLKAVDE 97

Db 605 KIEE-----LDEKIKAEAGKGDYKEAEKLEKAKLEKELKKLEQE 645

RESULT 12

F75216

hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)

C;Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C;Accession: F75216

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A;Reference number: A75001

A;Accession: F75216

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-281 <RAW>

A;Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49181

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB2181

Query Match 21.2%; Score 103; DB 2; Length 281;

Best Local Similarity 27.6%; Pred. No. 2.9;

Matches 34; Conservative 29; Mismatches 32; Indels 28; Gaps 5;

Qy 1 LKE-IDSDSDYLKGERAPLOSGLDTTKAKL-----SKLEELSDKIDELDA 47

Db 120 IKEVWAREEYKLLKEVEK--LKQEPFEEVKAKTEAAAELESLEKAKKEIEELKGVKEKLEQ 177

Qy 48 EIAKLEVLQKDAEGN-----NNVEAYFKGLEKTTAKKAELEKAEADLKAVD 96

Db 178 EKKEELKLEKSESVKLMVEYAKAKRAAELEAKLREYEEKVREE--ELERKVSLESLN 235

Qy 97 EPE 99

Db 236 EYE 238

RESULT 13

A28313

glued protein - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004

C;Accession: A28313

R;Swaroop, A.; Swaroop, M.; Garen, A.

Proc. Natl. Acad. Sci. U.S.A. 84, 6501-6505, 1987

A;Title: Sequence analysis of the complete cDNA and encoded polypeptide for the glued

A;Reference number: A28313; MUID:87317680; PMID:2819881

A;Accession: A28313

A;Molecule type: DNA; mRNA

A;Residues: 1-1319 <SWA>

A;Cross-references: UNIPROT:P13496

A;Note: the authors' translation is inconsistent with the nucleotide sequence in the reg

C;Genetics:

A;Gene: FlyBase:Gl

A;Cross-references: FlyBase:FBgn0001108

A;Introns: 18/2; 479/3

C;Keywords: cytoskeleton; glycoprotein

F;397,590,771,888,980,1110,1127,1133,1142/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 21.2%; Score 103; DB 2; Length 1319;

Best Local Similarity 32.7%; Pred. No. 13;

Matches 33; Conservative 20; Mismatches 26; Indels 22; Gaps 4;

Qy 1 LKSIDSDSDYLKGERAPLOSGLDTTKAKLSKL---EELSDKIDELDAETAKLEVLQ 56

Db 429 LRDLASHDKHDIQK-----LSKELEMKRSEVTELETRTEKLSAKIDELAIADLQEQV 482

Qy 57 KDAEGNNVNEAYFKGLEKTTAKKAELEKAEADLKAVDE 97

Db 483 DAALG-----AEWVEQLAEKKMELE-----DKVKLLDE 511

RESULT 14

A26655

myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)  
N:Contains: myosin Arpase (EC 3.6.4.1)  
C:Species: Dictyostelium discoideum  
C>Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 09-Jul-2004  
C:Accession: A26655; A24728; S00250  
R:Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.  
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986  
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum  
A:Reference number: A26655; PMID:87092266; PMID:3540939  
A:Accession: A26655  
A:Molecule type: DNA  
A:Residues: 1-2116 <WAR>  
A:Cross-references: UNIPROT:P08799; GB:M14628; GB:M11938; NID:9167834; PIDN:AAA33227.1; R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985  
A:Reference number: A24728; PMID:86016788; PMID:3901008  
A:Accession: A24728  
A:Molecule type: mRNA  
A:Residues: 2035-2116 <DEL>  
R:Wagie, G.; Noegel, A.; Scheel, J.; Gerisch, G.  
FEBS Lett. 227, 71-75, 1988  
A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium A:Reference number: S00250; PMID:88112256; PMID:2828113  
A:Accession: S00250  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1734-1893 <WAG>  
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphate binding site  
F:1-818/Domain: globular head <HED>  
F:89-747/Domain: myosin motor domain homology <MMOT>  
F:179-186/Region: nucleotide-binding motif A (P-loop)  
F:819-2116/Domain: alpha-helical rod <ROD>

Query Match 21.2%; Score 103; DB 2; Length 2116;  
Best Local Similarity 29.2%; Pred. No. 21;  
Matches 38; Conservative 17; Mismatches 43; Indels 32; Gaps 5;  
QY 1 LKEIDSDSE--DYLKEGERAPLQSKLDTKAKLKLSELSKID-----E 44  
DB 1453 IKRLNEELSELSVLEADER-CNSAIKAKTAESLSEKDEIDANNKAKAEKSK 1511  
QY 45 LDAETIAKLVQLKDAEGNNVVEAYFKEG-----LEKTTA-----EKKALEKAE 89  
DB 1512 LEVRVAELSESLDGKGTGVNVEFIRKDAEIDDLRLDRETESRIKSDKDKNTRKQFA 1571  
QY 90 DLKKAVIDEPE 99  
DB 1572 DLEAKVEEAQ 1581

RESULT 15  
AII333  
ABC transporter (ATP-binding protein) homolog lmo2073 [imported] - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AII333  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simeas, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species  
A:Reference number: AB1077; PMID:21537279; PMID:11679669  
A:Accession: AII333  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-650 <GLA>  
A:Cross-references: UNIPROT:Q8Y519; GB:NC\_003210; PIDN:CAD00151.1; PID:g16411543; GSPDB: A:Experimental source: strain EGD-e  
C:Genetics:

A:Gene: lmo2073  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
Query Match 21.1%; Score 102.5; DB 2; Length 650;  
Best Local Similarity 31.8%; Pred. No. 7.2;  
Matches 35; Conservative 22; Mismatches 32; Indels 21; Gaps 4;  
QY 2 KEIDSDSDYDKEGERRAPLQS---KLDTKAK-----LSKLELSKIDELDAEI 49  
DB 541 KELARLDAEDRRKGEQVEATASVRKLNQEEKEQQLRQRKRKLEEEKSMEETDEKI 600  
QY 50 AKLEVQLKDAEGNNVVEAYFKEGLEKTTAEKKALEKAEADLKKAVIDEPE 99  
DB 601 AELEQLTNPE-----VFQDHEKALEIT-----QELDAVRADGKGLMEWE 641

Search completed: June 21, 2005, 10:12:02  
Job time : 10.9 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 61.3194 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-16  
Perfect score: 485  
Sequence: 1 LKEIDSESDYLKEGERAP.....KKAELEKARADLKKAVIDEPE 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID       | Description        |
|------------|-------|---------------|--------|----------|--------------------|
| 1          | 477   | 98.4          | 739    | 2 Q9RQT4 | Q9RQT4 streptococc |
| 2          | 477   | 98.4          | 820    | 2 Q9RQT1 | Q9RQT1 streptococc |
| 3          | 477   | 98.4          | 929    | 2 Q9KK19 | Q9KK19 streptococc |
| 4          | 477   | 98.4          | 929    | 2 Q9ZAY5 | Q9ZAY5 streptococc |
| 5          | 458   | 94.4          | 437    | 2 Q9LAY4 | Q9LAY4 streptococc |
| 6          | 450   | 92.8          | 249    | 2 Q9L575 | Q9L575 streptococc |
| 7          | 444   | 91.5          | 99     | 2 Q8KK4  | Q8KK4 streptococc  |
| 8          | 441   | 90.9          | 224    | 2 Q8GNS8 | Q8GNS8 streptococc |
| 9          | 441   | 90.9          | 426    | 2 Q9LAY5 | Q9LAY5 streptococc |
| 10         | 437   | 90.1          | 395    | 2 Q9LAY2 | Q9LAY2 streptococc |
| 11         | 437   | 90.1          | 408    | 2 Q9LAY0 | Q9LAY0 streptococc |
| 12         | 433.5 | 89.4          | 869    | 2 Q9KX27 | Q9KX27 streptococc |
| 13         | 413   | 85.2          | 619    | 2 Q54972 | Q54972 streptococc |
| 14         | 413   | 85.2          | 619    | 2 Q8DR10 | Q8DR10 streptococc |
| 15         | 411   | 84.7          | 417    | 2 Q9LAY3 | Q9LAY3 streptococc |
| 16         | 391   | 80.6          | 415    | 2 Q9LAY1 | Q9LAY1 streptococc |
| 17         | 330.5 | 68.1          | 222    | 2 Q9L577 | Q9L577 streptococc |
| 18         | 330.5 | 68.1          | 225    | 2 Q9L591 | Q9L591 streptococc |
| 19         | 330.5 | 68.1          | 262    | 2 Q9L576 | Q9L576 streptococc |
| 20         | 330.5 | 68.1          | 415    | 2 Q9LAY7 | Q9LAY7 streptococc |
| 21         | 329.5 | 67.9          | 394    | 2 Q9LAY6 | Q9LAY6 streptococc |
| 22         | 329.5 | 67.9          | 395    | 2 Q9LAZ1 | Q9LAZ1 streptococc |
| 23         | 327.5 | 67.5          | 194    | 2 Q9L5B5 | Q9L5B5 streptococc |
| 24         | 327.5 | 67.5          | 218    | 2 Q6UBH2 | Q6UBH2 streptococc |
| 25         | 327.5 | 67.5          | 233    | 2 Q9L568 | Q9L568 streptococc |
| 26         | 327.5 | 67.5          | 236    | 2 Q9L569 | Q9L569 streptococc |
| 27         | 327.5 | 67.5          | 243    | 2 Q9L564 | Q9L564 streptococc |
| 28         | 327.5 | 67.5          | 243    | 2 Q9L567 | Q9L567 streptococc |
| 29         | 327.5 | 67.5          | 244    | 2 Q9L565 | Q9L565 streptococc |
| 30         | 327.5 | 67.5          | 246    | 2 Q9L578 | Q9L578 streptococc |
| 31         | 327.5 | 67.5          | 247    | 2 Q9L566 | Q9L566 streptococc |

|    |       |      |     |          |                    |
|----|-------|------|-----|----------|--------------------|
| 32 | 327.5 | 67.5 | 249 | 2 Q9L570 | Q9L570 streptococc |
| 33 | 327.5 | 67.5 | 254 | 2 Q9L563 | Q9L563 streptococc |
| 34 | 327.5 | 67.5 | 401 | 2 Q9LAZ2 | Q9LAZ2 streptococc |
| 35 | 326.5 | 67.3 | 255 | 2 Q9L581 | Q9L581 streptococc |
| 36 | 326.5 | 67.3 | 255 | 2 Q9L5B6 | Q9L5B6 streptococc |
| 37 | 323.5 | 66.7 | 416 | 2 Q9LAY8 | Q9LAY8 streptococc |
| 38 | 319.5 | 65.9 | 393 | 2 Q9LAZ3 | Q9LAZ3 streptococc |
| 39 | 316.5 | 65.3 | 406 | 2 Q9LAZ0 | Q9LAZ0 streptococc |
| 40 | 315.5 | 65.1 | 340 | 2 Q8KK5  | Q8KK5 streptococc  |
| 41 | 310.5 | 64.0 | 237 | 2 Q9L592 | Q9L592 streptococc |
| 42 | 310.5 | 64.0 | 395 | 2 Q9LAY9 | Q9LAY9 streptococc |
| 43 | 301.5 | 62.2 | 207 | 2 Q8GNS9 | Q8GNS9 streptococc |
| 44 | 195.5 | 40.3 | 653 | 2 Q34097 | Q34097 streptococc |
| 45 | 174.5 | 36.0 | 246 | 2 Q9L5B4 | Q9L5B4 streptococc |

ALIGNMENTS

RESULT 1  
Q9RQT4 ID Q9RQT4 PRELIMINARY; PRT; 739 AA.  
AC Q9RQT4;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Hypothetical protein pspC (fragment).  
GN Name=pspC;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E134;  
RX MEDLINE=20038319; PubMed=10569772;  
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;  
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic  
RT protein, PspC, which elicits cross-reactive antibodies to PspA and  
RT provides immunity to pneumococcal bacteremia.";  
RL Infect. Immun. 67:6533-6542(1999).  
DR EMBL; AF068647; AAF13457.1;  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR002479; CW\_binding.  
DR InterPro; IPR005873; Gpos\_Y5IRK.  
DR InterPro; IPR009053; Prefoldin.  
DR InterPro; IPR007756; RICH.  
DR Pfam; PF01473; CW\_binding\_1; 1.  
DR Pfam; PF05062; RICH; 2.  
DR Pfam; PF04650; Y5IRK\_signal; 1.  
DR TIGRFAM; TIGR01168; Y5IRK\_signal; 1.  
KW Hypothetical protein.  
FT NON\_TER 739  
SQ SEQUENCE 739 AA; 33960 MW; 7EE2F2F676ABF989 CRC64;  
  
Query Match 98.4%; Score 477; DB 2; Length 739;  
Best Local Similarity 99.0%; Pred. No. 2.4e-22;  
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 LKEIDSESDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60  
|||||  
DB 537 LKEIDSESDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 596  
|||||  
QY 61 GNNVVEYFKGLEKTTAEKKALEKAEADLKKAVIDEPE 99  
|||||  
DB 597 GNNVVEYFKGLEKTTAEKKALEKAEADLKKAVIDEPE 635  
|||||  
  
RESULT 2  
Q9RQT1 ID Q9RQT1 PRELIMINARY; PRT; 820 AA.  
AC Q9RQT1;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TPR 820 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 98.4%; Score 477; DB 2; Length 820;
Best Local Similarity 99.0%; Pred. No. 2.6e-22;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 60
Db 530 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 589

Qy 61 GNNNVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 628

RESULT 3
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=srcf10;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli P., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.

```

```

DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC82933302FAFA64 CRC64;

Query Match 98.4%; Score 477; DB 2; Length 929;
Best Local Similarity 99.0%; Pred. No. 2.9e-22;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 60
Db 530 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 589

Qy 61 GNNNVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 628

RESULT 4
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC82933302FFB081 CRC64;

Query Match 98.4%; Score 477; DB 2; Length 929;
Best Local Similarity 99.0%; Pred. No. 2.9e-22;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 60
Db 530 LKEIDESDSEDLKGEGRAPLQSKLDTKKAQKLELSKIDELDAEIAKLEVLQKDAE 589

Qy 61 GNNNVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 99
Db 590 GNNNVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 628

RESULT 5
Q9LAY4 ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```



```
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.B.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae."
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PF01473; CW_binding_1; 1.
FT NON_TER 437 437
SQ SEQUENCE 437 AA; 48071 MW; 66BFD2CD13E08CD8 CRC64;

Query Match 94.4%; Score 458; DB 2; Length 437;
Best Local Similarity 96.0%; Pred. No. 2.3e-21;
Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGEERAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 235 LKEIDSDSDYKGEERAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 294
QY 61 GNNVEAYFKEGLEKTTAEKKALEKAEADLKKAVIDE 99
DB 295 GNNVEAYFKEGLEKTTAEKKALEKAEADLKKAVIDE 333

RESULT 6
Q9L575 PRELIMINARY; PRT; 249 AA.
ID AC Q9L575;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones."
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 1 1
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 92.8%; Score 450; DB 2; Length 249;
Best Local Similarity 92.9%; Pred. No. 4.5e-21;
Matches 92; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGEERAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 1 LKEIDSDSDYKGEERAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60

RESULT 7
Q8KQK4 PRELIMINARY; PRT; 99 AA.
ID AC Q8KQK4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
immunization with DNA vaccines against Streptococcus pneumoniae
expressing PspA fragments from different clades."
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082388; AAL92493.1; -.
FT NON_TER 1 1
FT NON_TER 99 99
SQ SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match 91.5%; Score 444; DB 2; Length 99;
Best Local Similarity 91.9%; Pred. No. 4.6e-21;
Matches 91; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYKGEERAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
DB 1 LKEIDSDSDYKGEERAPLQSKLDTKAKLSKLEELSDKIDELDAETAKLEVQLKDAE 60
QY 61 GNNVEAYFKEGLEKTTAEKKALEKAEADLKKAVIDE 99
DB 61 GNNVEAYFKEGLEKTTAEKKALEKAEADLKKAVIDE 99

RESULT 8
Q8GNS8 PRELIMINARY; PRT; 224 AA.
ID AC Q8GNS8;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EN124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
Lorino G., Recchia S., Fantosti A., Beall B.;
RT "Genotypes of invasive pneumococcal isolates recently recovered from
Italian patients."
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
FT HSP; P00192; IAPC.
```

```

DR InterPro; IPR009082; his_kin_homodim.
FT NON TER 1 224
RC STRAIN=224 AA; 23418 MW; 48674E27AFB66A95 CRC64;
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

Query Match 90.9%; Score 441; DB 2; Length 224;
Best Local Similarity 90.9%; Pred. No. 1.5e-20;
Matches 90; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 17 LKIDNESDSDYVKEGFRAPLQSELDTKAKLLKLELSKIEELDAEIAEVLKDAE 76
Qy 61 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 99
Db 77 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 115

RESULT 9
Q9LAY5 PRELIMINARY; PRT; 426 AA.
AC Q9LAY5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match 90.9%; Score 441; DB 2; Length 426;
Best Local Similarity 90.9%; Pred. No. 2.7e-20;
Matches 90; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 215 LKIDNESDSDYVKEGFRAPLQSELDTKAKLLKLELSKIEELDAEIAEVLKDAE 274
Qy 61 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 99
Db 275 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 313

RESULT 10
Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```

---

```

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 395
SQ SEQUENCE 395 AA; 42963 MW; 58E6EF956BCBCC1E CRC64;

Query Match 90.1%; Score 437; DB 2; Length 395;
Best Local Similarity 90.9%; Pred. No. 4.5e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 225 LKEINESDSDYAKGFRAPLQSKLDAKAKLLKLELSKIEELDAEIAEVLKDAE 284
Qy 61 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 99
Db 285 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 323

RESULT 11
Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON TER 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 90.1%; Score 437; DB 2; Length 408;
Best Local Similarity 90.9%; Pred. No. 4.6e-20;
Matches 90; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKEIDSDSDYLKEGERAPLQSKLDTKAKLSKLELSKIDELDAEIAKLEVLKDAE 60
Db 228 LKEINESDSDYAKGFRAPLQSKLDAKAKLLKLELSKIEELDAEIAEVLKDAE 287
Qy 61 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 99
Db 288 GNNVVEAYFKEGLEKTTAAEKAELKAEADLKKAVDEPE 326

RESULT 12

```

```
Q9KK27
ID Q9KK27 PRELIMINARY; PRT; 869 AA.
AC Q9KK27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=95;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSSP; P06653; 1H8G.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
DR SEQUENCE 869 AA; 98732 MW; AFF2B504347E0220 CRC64;

Query Match 89.4%; Score 433.5; DB 2; Length 869;
Best Local Similarity 91.9%; Pred. No. 1.5e-19;
Matches 91; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLVQLKDAE 60
Db 537 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDVN-CNLRSQLKDAE 595
QY 61 GNNVVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 99
Db 596 GNNVVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 634

RESULT 13
Q54972 PRELIMINARY; PRT; 619 AA.
ID Q54972;
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
SEQUENCE FROM N.A.
RC MEDLINE=92105030; PubMed=17292249;
RA Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis.";
RL J. Bacteriol. 174:601-609(1992).
RN [2]
SEQUENCE FROM N.A.
RA Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.

HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW SIGNAL.
FT SIGNAL. 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 85.2%; Score 413; DB 2; Length 619;
Best Local Similarity 87.9%; Pred. No. 2.2e-18;
Matches 87; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLVQLKDAE 60
Db 223 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLVQLKDAE 282
QY 61 GNNVVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 99
Db 283 ENNVEDYFKEGLEKTTAAKKALEKAEADLKKAVNEPE 321

RESULT 14
Q8DR10 PRELIMINARY; PRT; 619 AA.
ID Q8DR10;
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name=pspA; OrderedLocNames=spr0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
SEQUENCE FROM N.A.
RC MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Complete proteome.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 85.2%; Score 413; DB 2; Length 619;
Best Local Similarity 87.9%; Pred. No. 2.2e-18;
Matches 87; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLVQLKDAE 60
Db 223 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLVQLKDAE 282
QY 61 GNNVVEAYFKEGLEKTTAAKKALEKAEADLKKAVDEPE 99
Db 283 ENNVEDYFKEGLEKTTAAKKALEKAEADLKKAVNEPE 321
```

```
RESULT 15
Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ef10197;
RX MEDLINE=2048953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
  in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSSP; P00192; 256B.
FT NON TER 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match      84.7%; Score 411; DB 2; Length 417;
Best Local Similarity 86.9%; Pred. No. 2e-18;
Matches 86; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy      1 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 60
Db      213 LKEIDSDSDYLKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVLQKDAE 272

Qy      61 GNNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db      273 ENNVVEDYFKEGLEKTTAAKKAELKAEADLKKAVNEPE 311
```

Search completed: June 21, 2005, 10:22:12  
Job time : 61.3194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 73.8459 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PKRINLSQVXLKVCRAP.....KKAELXAXADLKXAVDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp19808:\*  
2: Geneseqp19908:\*  
3: Geneseqp20008:\*  
4: Geneseqp20018:\*  
5: Geneseqp20028:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID         | Description |
|------------|-------|-------------|--------|------------|-------------|
| 1          | 453   | 97.4        | 190    | 2 AAW14569 | Streptococ  |
| 2          | 453   | 97.4        | 193    | 7 ABW02603 | Bg9163c p   |
| 3          | 453   | 97.4        | 8991   | 6 ABU08487 | S. pneumo   |
| 4          | 361   | 77.6        | 206    | 2 AAW14574 | Streptoco   |
| 5          | 361   | 77.6        | 206    | 7 ABW02608 | Db15c pne   |
| 6          | 355   | 76.3        | 170    | 7 ABW02614 | Rct135c p   |
| 7          | 355   | 76.3        | 181    | 7 ABW02596 | 0922134c    |
| 8          | 355   | 76.3        | 865    | 6 ABU08489 | S. pneumo   |
| 9          | 355   | 76.3        | 929    | 2 AAW14593 | Streptoco   |
| 10         | 355   | 76.3        | 929    | 2 AAY43384 | S. pneumo   |
| 11         | 354   | 76.1        | 204    | 2 AAW14578 | Streptoco   |
| 12         | 354   | 76.1        | 204    | 7 ABW02612 | p           |
| 13         | 352   | 75.7        | 188    | 2 AAW14580 | Streptoco   |
| 14         | 352   | 75.7        | 188    | 7 ABW02613 | p           |
| 15         | 347   | 74.6        | 1231   | 6 ABU08490 | Fragment    |
| 16         | 343   | 73.8        | 588    | 6 ABU08491 | Coiled co   |
| 17         | 343   | 73.8        | 589    | 2 AAY43392 | PapC alph   |
| 18         | 340.5 | 73.2        | 180    | 2 AAW14562 | Streptoco   |
| 19         | 337.5 | 72.6        | 187    | 2 AAW14579 | Streptoco   |
| 20         | 312   | 67.1        | 605    | 6 ABU08493 | Fragment    |
| 21         | 307   | 66.0        | 198    | 2 AAW14581 | Streptoco   |
| 22         | 304   | 65.4        | 198    | 7 ABW02615 | Rx1c pneu   |
| 23         | 304   | 65.4        | 315    | 2 AAY04375 | Streptoco   |
| 24         | 304   | 65.4        | 619    | 2 AAR87598 | Pneumococ   |
| 25         | 304   | 65.4        | 619    | 2 AAR87598 | Pneumococ   |

|    |     |      |     |            |           |
|----|-----|------|-----|------------|-----------|
| 26 | 304 | 65.4 | 619 | 2 AAR86911 | Pneumococ |
| 27 | 304 | 65.4 | 619 | 2 AAY41838 | Streptoco |
| 28 | 304 | 65.4 | 619 | 5 AAE18782 | S. pneumo |
| 29 | 304 | 65.4 | 619 | 6 ABU45778 | Protein e |
| 30 | 304 | 65.4 | 619 | 8 ADO52126 | Streptoco |
| 31 | 304 | 65.4 | 648 | 2 AAW70336 | Pneumococ |
| 32 | 304 | 65.4 | 648 | 2 AAW62274 | Streptoco |
| 33 | 304 | 65.4 | 648 | 2 AAY41837 | Streptoco |
| 34 | 304 | 65.4 | 648 | 2 AAW87879 | A pneumoc |
| 35 | 304 | 65.4 | 653 | 2 AAW92456 | S. pneumo |
| 36 | 304 | 65.4 | 684 | 2 AAR73912 | Streptoco |
| 37 | 297 | 63.9 | 204 | 2 AAW14571 | Streptoco |
| 38 | 297 | 63.9 | 204 | 7 ABW02605 | Ef1019c p |
| 39 | 294 | 63.2 | 289 | 2 AAW62276 | Streptoco |
| 40 | 294 | 63.2 | 289 | 2 AAY41840 | Streptoco |
| 41 | 294 | 63.2 | 289 | 2 AAW87910 | Protein s |
| 42 | 294 | 63.2 | 289 | 2 AAW92458 | S. pneumo |
| 43 | 293 | 63.0 | 653 | 2 AAR27150 | PspA frag |
| 44 | 277 | 59.6 | 195 | 2 AAW14591 | Streptoco |
| 45 | 277 | 59.6 | 195 | 7 ABW02625 | Wuzc pneu |

#### ALIGNMENTS

RESULT 1

AAW14569

ID AAW14569 standard; protein; 190 AA.

XX AAW14569;

XX AC AAW14569;

XX 17-OCT-2003 (revised)

DT 28-OCT-1997 (first entry)

XX XX

Streptococcus pneumoniae PspA central region.

DE XX

PspA; pneumococcal surface protein; vaccine; otitis media; meningitis; bacteraemia; pneumonia.

KW XX

Streptococcus pneumoniae; strain Bg8743.

XX XX

Key Location/Qualifiers

FT Misc-difference 21 /note= "unidentified amino acid"

FT Misc-difference 24 /note= "unidentified amino acid"

FT Misc-difference 95 /note= "unidentified amino acid"

FT Misc-difference 97 /note= "unidentified amino acid"

FT Misc-difference 186 /note= "unidentified amino acid"

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

XX Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

Streptococcus pneumoniae surface protein PspC and truncated PspA - used in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.



```

CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 206 AA;
    Query Match      77.6%; Score 361; DB 2; Length 206;
    Best Local Similarity 90.2%; Pred. No. 2.8e-31;
    Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0
QY 18 RAPLQSKLDQAQKAEELKLEELSGKIKELDAIEAEVQLKDAEGNNNVEAFYFKEGLEKTT 77
    |||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||
Db 18 RAPLQSELDTKKAKLLKLEELSGKIEELDAIEAEVQLKDAEGNNNVEAFYFKEGLEKTT 77
    QY 78 AEKKAELXXXADLKKAVDEPE 99
    |||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||
Db 78 AEKKAELKAEADLKKAVDEPE 99

RESULT 5
ABW02608
ID ABW02608 standard; protein; 206 AA.
XX
AC ABW02608;
XX
DT 12-FEB-2004 (first entry)
XX
DB Db15c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..206
FT /note= "Xaa = Unknown amino acid"
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PP 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048996.
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
WI; 2003-862841/80.
XX
PT Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
PS Example 6; SEQ ID NO 54; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Dbl5c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX

```

SQ Sequence 206 AA;  
 Query Match 77.6%; Score 361; DB 7; Length 206;  
 Best Local Similarity 90.2%; Pred. No. 2.8e-31;  
 Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 RAPIQSKLDAQKAELELLKLEELSGKIKELDAIELEVLQKDAEGNNVAYFKEGLEKTT 77  
 DB 18 RAPIQSELDTKKAKLLKLEELSGKIELEDAIELEVLQKDAEGNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVIDEPE 99  
 DB 78 AEKKAEELEKAEADLKKAVIDEPE 99

RESULT 6  
 ABW02614  
 ID ABW02614 standard; protein; 170 AA.  
 XX  
 AC ABW02614;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Rct135c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX  
 DR WPI; 2003-862841/80.  
 XX  
 PT Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 PS Example 6; SEQ ID NO 60; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Rct135c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX

SQ Sequence 170 AA;  
 Query Match 76.3%; Score 355; DB 7; Length 170;  
 Best Local Similarity 78.6%; Pred. No. 9.8e-31;  
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVKLXXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAIELEVLQKDAEG 61  
 DB 2 KEIDESDSEDLKXGLRAPLQSKLDTKAKKLKLEELSDKIDELDAIEIAKLEVLQKDAEG 61

QY 62 NNNVEAYFKEGLEKTTAEKKAEELEXAXADLKKAVIDEPE 99  
 DB 62 NNNVEAYFKEGLEKTTAEKKAEELEKAEADLKKAVIDEPE 99

RESULT 7  
 ABW02596  
 ID ABW02596 standard; protein; 181 AA.  
 XX  
 AC ABW02596;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE 0922134c pneumococcal surface protein A (PspA) central region.  
 XX  
 KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KW immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.  
 XX  
 PN US6592876-B1.  
 XX  
 PD 15-JUL-2003.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
 XX  
 DR WPI; 2003-862841/80.  
 XX  
 PT Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprises at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX  
 PS Example 6; SEQ ID NO 42; 121pp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspAs) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is 0922134c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX

SQ Sequence 181 AA;  
 Query Match 76.3%; Score 355; DB 7; Length 181;  
 Best Local Similarity 78.6%; Pred. No. 1.1e-30;  
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVKLXXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAIELEVLQKDAEG 61  
 DB 2 KEIDESDSEDLKXGLRAPLQSKLDTKAKKLKLEELSDKIDELDAIEIAKLEVLQKDAEG 61

QY 62 NNNVEAYFKEGLEKTTAEKKAEELEXAXADLKKAVIDEPE 99  
 DB 62 NNNVEAYFKEGLEKTTAEKKAEELEKAEADLKKAVIDEPE 99





Query Match 76.3%; Score 355; DB 2; Length 929;  
 Best Local Similarity 78.6%; Pred. No. 8.9e-30;  
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAEILLKLEELSGKIKELDAIEAEVQLKDAEG 61  
 DB 531 KEIDESDSEYDKGLRAPLQSKLDTTKKAKLSKLEELSDKIDELDAIEAEVQLKDAEG 590  
 QY 62 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 99  
 DB 591 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 628

## RESULT 10

AAV43384

ID AAY43384 standard; protein; 929 AA.

XX AC AAY43384;

XX DT 27-JAN-2000 (first entry)

XX DE S. pneumoniae PspC protein sequence.

XX KW PspC gene; pneumococcal surface protein C; epitope; diagnosis;  
 XX KW epitopic region; immunogenic composition; vaccine composition; therapy;  
 XX KW meningitis; immunological response; otitis media; bacteraemia; pneumonia;  
 XX KW anti-PspA antibody.

XX OS Streptococcus pneumoniae.

XX PN W09953940-A1.

XX PD 28-OCT-1999.

XX PF 23-APR-1999; 99WO-US008895.

XX PR 23-APR-1998; 98US-0082728P.

XX PA (UYAL-) UNIV ALABAMA.

XX PI Briles DE, Hollingshead SK, Brooks-Walter A;

XX DR WPI; 1999-620581/53.

XX DR N-PSDB; AAZ31956.

XX FT New epitope polypeptides of Pneumococcal surface protein C, used to  
 develop products for immunological, immunogenic or vaccine compositions,  
 particularly against Streptococcus pneumoniae infections.

XX PS Example; Fig 11; 109pp; English.

XX CC This sequence is the Streptococcus pneumoniae pneumococcal surface  
 CC protein C. The invention relates to an isolated and/or purified  
 CC polypeptide comprising at least one epitope or epitopic region of  
 CC Pneumococcal surface protein C (PspC). The polypeptides or vectors  
 CC containing sequence encoding them can be used as immunogenic,  
 CC immunological or vaccine compositions. The compositions can be used for  
 CC eliciting an immunological response against Streptococcus pneumoniae  
 CC (SP), which can cause otitis media, meningitis, bacteraemia and  
 CC pneumonia. They can be used for eliciting an anti-PspA antibody. The  
 CC nucleic acid molecules can also be used for detecting pspC, pspA or SP

XX SQ Sequence 929 AA;

Query Match 76.3%; Score 355; DB 2; Length 929;  
 Best Local Similarity 78.6%; Pred. No. 8.9e-30;  
 Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAEILLKLEELSGKIKELDAIEAEVQLKDAEG 61  
 DB 531 KEIDESDSEYDKGLRAPLQSKLDTTKKAKLSKLEELSDKIDELDAIEAEVQLKDAEG 590

QY 62 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 99  
 DB 591 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDEPE 628

## RESULT 11

AAW14578

ID AAW14578 standard; protein; 204 AA.

XX AC AAW14578;

XX DT 17-OCT-2003 (revised)

XX DT 28-OCT-1997 (first entry)

XX DE Streptococcus pneumoniae PspA central region.

XX KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 XX KW bacteraemia; pneumonia.

XX OS Streptococcus pneumoniae; strain Rct123.

XX FH Key Location/Qualifiers

FT Misc-difference 4 /note= "unidentified amino acid"

FT Misc-difference 8 /note= "unidentified amino acid"

XX PN W09709994-A1.

XX PD 20-MAR-1997.

XX PF 16-SEP-1996; 96WO-US014819.

XX PR 15-SEP-1995; 95US-00529055.

XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;

XX PI Hollingshead S, Tart R, Brooks-Walter A;

XX DR WPI; 1997-202002/18.

XX ST Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 in vaccines for protecting animals against S.pneumoniae infection.

XX PS Example 6; Fig 13; 296pp; English.

XX CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Rct123.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 204 AA;

Query Match 76.1%; Score 354; DB 2; Length 204;  
 Best Local Similarity 86.2%; Pred. No. 1.6e-30;  
 Matches 75; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 13 LKXVCRAPIQSKLDAQKAEILLKLEELSGKIKELDAIEAEVQLKDAEGNNNVEAYFKEG 72  
 DB 13 LKGLRAPLQSKLDTTKKAKLSKLEELSDKIDELDAIEAEVQLKDAEGNNNVEAYFKEG 72  
 QY 73 LEKTTAEKKAELXAXADLKKAVDEPE 99

Db 73 LEKTTAEKKAEELEKAEADLKKANDEPE 99

RESULT 12  
ABW02612

ID ABW02612 standard; protein; 204 AA.

AC ABW02612;

XX 12-FEB-2004 (first entry)

XX Rct123c pneumococcal surface protein A (PspA) central region.

XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

XX immunological; gene therapy; immunostimulant.

XX Unidentified.

XX OS

XX Key Location/Qualifiers

PH Misc-difference 1..204

FT /note= "Xaa = Unknown amino acid"

FT US6592876-B1.

PN 15-JUL-2003.

PD

XX 15-SEP-1995; 95US-00529055.

PF 20-APR-1993; 93US-00048896.

PR 06-JUN-1995; 95US-00465746.

PR

XX (UABR-) UAB RES FOUND.

PA

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

PI WPI; 2003-862841/80.

XX

XX Immunological composition for obtaining expression products used for

PT detecting the presence of Streptococcus pneumoniae or its strain,

PT comprises at least two different full length isolated gene encoding

PT pneumococcal surface protein A.

XX

XX Example 6; SEQ ID NO 58; 121pp; English.

PS

XX The present invention relates to an immunological composition comprising

CC at least 2 different full length isolated genes encoding pneumococcal

CC surface protein A (PspAs) from different groups based on restriction

CC fragment polymorphism analysis. The invention is useful for obtaining

CC expression products by recombinant techniques to detect, determine,

CC isolate or diagnose the presence of Streptococcus pneumoniae or its

CC strain. The expression product is useful for preparing antigenic,

CC immunological or vaccine compositions, for eliciting antibodies, an

CC immunological response (other than or additional to antibodies) or a

CC protective response (including antibody or other immunological response

CC by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Rct123c

CC pneumococcal surface protein A (PspA) central region. This sequence is

CC used in the exemplification of the invention

XX

XX Sequence 204 AA;

Query Match 76.1%; Score 354; DB 7; Length 204;

Best Local Similarity 86.2%; Pred. No. 1.6e-30;

Matches 75; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 13 LKVCRAFLQSLKDAQAELILKLELSGKIKELDAETAELEVLQKDAEGNNVAYFKEG 72

Db 13 LKEGLRAPLQSLTKTKAKLSKLELSGKIDELDAETAELEVLQKDAEGNNVAYFKEG 72

Qy 73 LEKTTAEKKAEELEKAEADLKKANDEPE 99

Db 73 LEKTTAEKKAEELEKAEADLKKANDEPE 99



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 18.4867 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PKRIMLSQKVLKXVCRAP.....KKALEXAXADLKKADEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 457   | 98.3        | 99     | 4     | US-09-147-875A-17 |
| 2          | 453   | 97.4        | 193    | 4     | US-08-529-055-49  |
| 3          | 453   | 97.4        | 8931   | 4     | US-08-714-741-32  |
| 4          | 439   | 94.4        | 99     | 2     | US-08-710-749-16  |
| 5          | 373   | 80.2        | 99     | 4     | US-09-147-875A-15 |
| 6          | 361   | 77.6        | 99     | 4     | US-09-147-875A-14 |
| 7          | 361   | 77.6        | 206    | 4     | US-08-529-055-54  |
| 8          | 359   | 77.2        | 99     | 2     | US-08-710-749-14  |
| 9          | 355   | 76.3        | 170    | 4     | US-08-529-055-60  |
| 10         | 355   | 76.3        | 181    | 4     | US-08-529-055-42  |
| 11         | 355   | 76.3        | 864    | 4     | US-08-714-741-40  |
| 12         | 354   | 76.1        | 204    | 4     | US-08-529-055-58  |
| 13         | 352   | 75.7        | 99     | 4     | US-09-147-875A-16 |
| 14         | 352   | 75.7        | 188    | 4     | US-08-529-055-59  |
| 15         | 347   | 74.6        | 99     | 2     | US-08-710-749-13  |
| 16         | 347   | 74.6        | 1231   | 4     | US-08-714-741-41  |
| 17         | 344   | 74.0        | 99     | 2     | US-08-710-749-17  |
| 18         | 343.5 | 73.9        | 100    | 4     | US-09-147-875A-10 |
| 19         | 343   | 73.8        | 141    | 4     | US-09-286-981B-2  |
| 20         | 343   | 73.8        | 588    | 4     | US-08-714-741-42  |
| 21         | 341   | 73.3        | 99     | 2     | US-08-710-749-15  |
| 22         | 312   | 67.1        | 605    | 4     | US-08-714-741-46  |
| 23         | 304   | 65.4        | 198    | 2     | US-08-710-749-11  |
| 24         | 304   | 65.4        | 198    | 4     | US-08-529-055-61  |
| 25         | 304   | 65.4        | 619    | 1     | US-08-465-746-2   |
| 26         | 304   | 65.4        | 619    | 1     | US-08-214-164-2   |
| 27         | 304   | 65.4        | 619    | 2     | US-08-467-852A-3  |

|    |     |      |     |   |                   |                    |
|----|-----|------|-----|---|-------------------|--------------------|
| 28 | 304 | 65.4 | 619 | 2 | US-08-246-636-2   | Sequence 2, Appli  |
| 29 | 304 | 65.4 | 619 | 2 | US-08-247-491A-3  | Sequence 2, Appli  |
| 30 | 304 | 65.4 | 619 | 2 | US-08-319-795-2   | Sequence 2, Appli  |
| 31 | 304 | 65.4 | 619 | 2 | US-08-468-985-2   | Sequence 2, Appli  |
| 32 | 304 | 65.4 | 619 | 3 | US-08-312-949-2   | Sequence 2, Appli  |
| 33 | 304 | 65.4 | 648 | 1 | US-08-072-070-2   | Sequence 2, Appli  |
| 34 | 304 | 65.4 | 648 | 1 | US-08-469-434-2   | Sequence 2, Appli  |
| 35 | 304 | 65.4 | 648 | 1 | US-08-214-222-2   | Sequence 2, Appli  |
| 36 | 304 | 65.4 | 648 | 2 | US-08-467-852A-2  | Sequence 2, Appli  |
| 37 | 304 | 65.4 | 648 | 2 | US-08-468-718-2   | Sequence 2, Appli  |
| 38 | 304 | 65.4 | 648 | 2 | US-08-247-491A-2  | Sequence 2, Appli  |
| 39 | 304 | 65.4 | 648 | 3 | US-08-446-201-3   | Sequence 3, Appli  |
| 40 | 304 | 65.4 | 695 | 1 | US-08-127-499A-23 | Sequence 23, Appli |
| 41 | 304 | 65.4 | 695 | 1 | US-08-482-847-23  | Sequence 23, Appli |
| 42 | 297 | 63.9 | 99  | 2 | US-08-710-749-10  | Sequence 10, Appli |
| 43 | 297 | 63.9 | 99  | 4 | US-09-147-875A-11 | Sequence 11, Appli |
| 44 | 297 | 63.9 | 204 | 4 | US-08-529-055-51  | Sequence 51, Appli |
| 45 | 294 | 63.2 | 289 | 1 | US-08-072-070-4   | Sequence 4, Appli  |

#### ALIGNMENTS

##### RESULT 1

US-09-147-875A-17  
; Sequence 17, Application US/09147875A  
; Patent No. 6638516  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(99)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-09-147-875A-17

Query Match 98.3%; Score 457; DB 4; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2e-42;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |   |   |    |
|----|---|---|----|
| QY | 1 | PKRIMLSQKVLKXVCRAPLOSLQKLDQAQKAEILKLEELSGKIKELDAETAEVQLKDAE | 60 |
| DB | 1 | PKRIMLSQKVLKXVCRAPLOSLQKLDQAQKAEILKLEELSGKIKELDAETAEVQLKDAE | 60 |

QY 61 GNNVAYPKGLEKTTAEKKALEAXADLKKADEPE 99

DB 61 GNNVAYPKGLEKTTAEKKALEAXADLKKADEPE 99

##### RESULT 2

US-08-529-055-49  
; Sequence 49, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yotter, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; TITLE OF INVENTION: Portions and Products  
; NUMBER OF SEQUENCES: 73

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-49

Query Match          97.4%; Score 453; DB 4; Length 193;
Best Local Similarity 99.0%; Pred. No. 1.2e-41;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKRIMSLSQVKVXKVCRAPIQSKLDQAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 60
Db 13 PKRIMSLSQVKVXKVCRAPIQSKLDQAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 72

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 73 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 111

RESULT 3
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match          97.4%; Score 453; DB 4; Length 8991;
Best Local Similarity 99.0%; Pred. No. 1.4e-39;
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKRIMSLSQVKVXKVCRAPIQSKLDQAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 60
Db 4958 PKRIMSLSQVKVXKVCRAPIQSKLDQAQKAEALLKLELSGKIKELDAEIAELEVLQKDAE 5017

Qy 61 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 5018 GNNVVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 5056

RESULT 4
US-08-710-749-16
; Sequence 16, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-16

Query Match      94.4%; Score 439; DB 2; Length 99;
Best Local Similarity 96.0%; Pred. No. 1.8e-40;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PKRMSLSQKXUKYVCRAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAE 60
DB 1 PKRMSLSQKXUKYVCRAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAE 60

QY 61 GNNVAYFKEGLEKTTAEKAELEXAXADLKKAVIDEPE 99
DB 61 GNNVAYFKEGLEKTTAEKAELEXAXADLKKAVIDEPE 99

RESULT 5
US-09-147-875A-15
; Sequence 15, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-15

Query Match      80.2%; Score 373; DB 4; Length 99;
Best Local Similarity 93.8%; Pred. No. 2.6e-33;
Matches 77; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 18 RAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAEAGNNNNVAYFKEGLEKTT 77
DB 18 RAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAEAGNNNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVIDEPE 99
DB 78 AEKKAEELEXAXADLKKAVIDEPE 99

RESULT 6
US-09-147-875A-14
; Sequence 14, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(99)
; OTHER INFORMATION: amino acid 'xaa' can be any amino acid
US-09-147-875A-14

Query Match      77.6%; Score 361; DB 4; Length 99;
Best Local Similarity 90.2%; Pred. No. 5.2e-32;
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 RAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAEAGNNNNVAYFKEGLEKTT 77
DB 18 RAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAEAGNNNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVIDEPE 99
DB 78 AEKKAEELEXAXADLKKAVIDEPE 99

US-08-529-055-54
; Sequence 54, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-54

Query Match      77.6%; Score 361; DB 4; Length 206;
Best Local Similarity 90.2%; Pred. No. 1.3e-31;
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 RAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAEAGNNNNVAYFKEGLEKTT 77
DB 18 RAPLQSKLDAQKAEELKLELSGKIKELDAETAELEVLQKDAEAGNNNNVAYFKEGLEKTT 77

QY 78 AEKKAEELEXAXADLKKAVIDEPE 99
DB 78 AEKKAEELEXAXADLKKAVIDEPE 99

RESULT 8
US-08-710-749-14
; Sequence 14, Application US/08710749
```

; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollinghead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-14

Query Match 77.2%; Score 359; DB 2; Length 99;  
Best Local Similarity 90.2%; Pred. No. 8.6e-32;  
Matches 74; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 18 RAPQSKLDKAKKLLKLEELSGKIKELDAEIAEVLQKDAEGNNVVEYFKGLEKTT 77  
Db 18 RAPQSKLDKAKKLLKLEELSGKIKELDAEIAEVLQKDAEGNNVVEYFKGLEKTT 77  
  
Qy 78 AEKKAELXAXADLKKAVIDEPE 99  
Db 78 AEKKAELXAXADLKKAVIDEPE 99  
  
RESULT 9  
US-08-529-055-60  
; Sequence 60, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY

; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2400  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 170 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-60  
  
Query Match 76.3%; Score 355; DB 4; Length 170;  
Best Local Similarity 78.8%; Pred. No. 4.6e-31;  
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;  
  
Qy 2 KRIMSLSQVXKXVCRAPIQSKLDQAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEG 61  
Db 2 KEIDESDSEYLAKEGLRAPLQSKLDTKAKKSKLEELSDKIDELDAEIAEVLQKDAEG 61  
  
Qy 62 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVIDEPE 99  
Db 62 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVIDEPE 99  
  
RESULT 10  
US-08-529-055-42  
; Sequence 42, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.



```
;
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-42

Query Match 76.3%; Score 355; DB 4; Length 181;
Best Local Similarity 78.6%; Pred. No. 4.9e-31;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAEIAEVLQKDAEG 61
Db 2 KEIDESSEYLYKEGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAEVLQKDAEG 61
QY 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99

RESULT 11
US-08-714-741-40
; Sequence 40, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 16-SEP-1996
; APPLICATION NUMBER: US/08/714,741
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Seq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

;
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-40

Query Match 76.3%; Score 355; DB 4; Length 864;
Best Local Similarity 78.6%; Pred. No. 3.4e-30;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLQKVLKXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAEIAEVLQKDAEG 61
Db 466 KEIDESSEYLYKEGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAEVLQKDAEG 525
QY 62 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 526 NNNVEAYFKEGLEKTTAEKKAELXAXADLKKAVIDEPE 563

RESULT 12
US-08-529-055-58
; Sequence 58, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-SEP-1995
; APPLICATION NUMBER: US/08/529,055
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-58

Query Match 76.1%; Score 354; DB 4; Length 204;
Best Local Similarity 86.2%; Pred. No. 7.4e-31;
Matches 75; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 13 LKXVCRAPIQSKLDAQKAELELLKLEELSGKIKELDAEIAEVLQKDAEGNNVYFKEG 72
Db 13 LKEGLRAPLQSKLDTTKAKLSKLEELSDKIDELDAEIAEVLQKDAEGNNVYFKEG 72
QY 73 LEKTTAEKKAELXAXADLKKAVIDEPE 99
Db 73 LEKTTAEKKAELXAXADLKKAVIDEPE 99
```

```
Db 73 LEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 13
US-09-147-875A-16
; Sequence 16, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147.875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 16
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-16

Query Match 75.7%; Score 352; DB 4; Length 99;
Best Local Similarity 78.6%; Pred. No. 5e-31;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRIMSLSQKVLKXVCRAPLQSKLDAQKAEALLKLEELSGKIKELDAIEAEVQLKDAEG 61
Db 2 KEIDESDSEYLYKEGERAPLQSKLDTKKAKLSKLEELSDKIDELDAIEAEVQLKDAEG 61

Qy 62 NNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 62 NNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99

RESULT 14
US-09-529-055-59
; Sequence 59, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529.055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 59:

Query Match 74.6%; Score 347; DB 2; Length 99;
Best Local Similarity 86.6%; Pred. No. 1.7e-30;
Matches 71; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 18 RAPLQSKLDAQKAEALLKLEELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKEGLEKTT 77
Db 18 RAPLQSELDTKKAKLLKLEELSGKIELEDAIEAEVQLKDVGGNNVVEAYFKEGLEKTT 77
Qy 78 AEKKAELKAEADLKKAVDEPE 99

; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-59

Query Match 75.7%; Score 352; DB 4; Length 188;
Best Local Similarity 77.6%; Pred. No. 1.1e-30;
Matches 76; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRIMSLSQKVLKXVCRAPLQSKLDAQKAEALLKLEELSGKIKELDAIEAEVQLKDAEG 61
Db 2 KEIDESDSEYLYKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAIEAEVQLKDAEG 61

Qy 62 NNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPE 99
Db 62 NNNVEAYFKEGLEKTTAEKKAELKAEADLKKAVDEPD 99

RESULT 15
US-08-710-749-13
; Sequence 13, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710.749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-13

Query Match 74.6%; Score 347; DB 2; Length 99;
Best Local Similarity 86.6%; Pred. No. 1.7e-30;
Matches 71; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 18 RAPLQSKLDAQKAEALLKLEELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKEGLEKTT 77
Db 18 RAPLQSELDTKKAKLLKLEELSGKIELEDAIEAEVQLKDVGGNNVVEAYFKEGLEKTT 77
Qy 78 AEKKAELKAEADLKKAVDEPE 99
```

Db           |||   |||   |   |||||||   |||  
78 AEMATELEKAEADLKXVDEPE 99

Search completed: June 21, 2005, 10:25:22  
Job time : 19.4867 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 63.2388 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PKRINSLSQKVLKXVCRAP.....KKAEELEXAXADLKXAVDEPE 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 457   | 98.3        | 99     | 15    | US-10-674-755-17  |
| 2          | 453   | 97.4        | 193    | 15    | US-10-299-636-64  |
| 3          | 373   | 80.2        | 99     | 15    | US-10-674-755-15  |
| 4          | 361   | 77.6        | 99     | 15    | US-10-674-755-14  |
| 5          | 361   | 77.6        | 206    | 15    | US-10-299-636-69  |
| 6          | 355   | 76.3        | 170    | 15    | US-10-299-636-75  |
| 7          | 355   | 76.3        | 181    | 15    | US-10-299-636-57  |
| 8          | 355   | 76.3        | 643    | 15    | US-10-299-636-95  |
| 9          | 355   | 76.3        | 670    | 9     | US-09-748-875-63  |
| 10         | 355   | 76.3        | 670    | 10    | US-09-298-523B-63 |
| 11         | 355   | 76.3        | 690    | 9     | US-09-748-875-61  |

|    |       |      |     |    |                      |                   |
|----|-------|------|-----|----|----------------------|-------------------|
| 12 | 355   | 76.3 | 690 | 10 | US-09-298-523B-61    | Sequence 61, Appl |
| 13 | 355   | 76.3 | 691 | 9  | US-09-748-875-1      | Sequence 1, Appl  |
| 14 | 355   | 76.3 | 691 | 10 | US-09-298-523B-1     | Sequence 1, Appl  |
| 15 | 355   | 76.3 | 701 | 9  | US-09-748-875-62     | Sequence 62, Appl |
| 16 | 355   | 76.3 | 701 | 10 | US-09-298-523B-62    | Sequence 62, Appl |
| 17 | 355   | 76.3 | 707 | 9  | US-09-748-875-2      | Sequence 2, Appl  |
| 18 | 355   | 76.3 | 707 | 10 | US-09-298-523B-2     | Sequence 2, Appl  |
| 19 | 355   | 76.3 | 711 | 9  | US-09-748-875-3      | Sequence 3, Appl  |
| 20 | 355   | 76.3 | 711 | 10 | US-09-298-523B-3     | Sequence 3, Appl  |
| 21 | 355   | 76.3 | 739 | 17 | US-10-732-923-3294   | Sequence 3294, Ap |
| 22 | 355   | 76.3 | 929 | 9  | US-09-748-875-60     | Sequence 60, Appl |
| 23 | 355   | 76.3 | 929 | 10 | US-09-298-523B-60    | Sequence 60, Appl |
| 24 | 355   | 76.3 | 929 | 15 | US-10-299-636-94     | Sequence 94, Appl |
| 25 | 354   | 76.1 | 204 | 15 | US-10-299-636-73     | Sequence 73, Appl |
| 26 | 352   | 75.7 | 99  | 15 | US-10-674-755-16     | Sequence 16, Appl |
| 27 | 352   | 75.7 | 188 | 15 | US-10-299-636-74     | Sequence 74, Appl |
| 28 | 343.5 | 73.9 | 100 | 15 | US-10-674-755-10     | Sequence 10, Appl |
| 29 | 343   | 73.8 | 141 | 14 | US-10-254-995-2      | Sequence 2, Appl  |
| 30 | 343   | 73.8 | 589 | 9  | US-09-748-875-14     | Sequence 14, Appl |
| 31 | 343   | 73.8 | 589 | 10 | US-09-298-523B-14    | Sequence 14, Appl |
| 32 | 343   | 73.8 | 589 | 15 | US-10-299-636-97     | Sequence 97, Appl |
| 33 | 326.5 | 70.2 | 336 | 15 | US-10-299-636-103    | Sequence 103, App |
| 34 | 304   | 65.4 | 198 | 15 | US-10-299-636-76     | Sequence 76, Appl |
| 35 | 304   | 65.4 | 354 | 15 | US-10-299-636-105    | Sequence 105, App |
| 36 | 304   | 65.4 | 588 | 15 | US-10-299-636-96     | Sequence 96, Appl |
| 37 | 304   | 65.4 | 619 | 10 | US-09-882-774-1      | Sequence 1, Appl  |
| 38 | 304   | 65.4 | 619 | 15 | US-10-282-122A-73702 | Sequence 73702, A |
| 39 | 304   | 65.4 | 619 | 16 | US-10-414-532-72     | Sequence 72, Appl |
| 40 | 297   | 63.9 | 99  | 15 | US-10-674-755-11     | Sequence 11, Appl |
| 41 | 297   | 63.9 | 204 | 15 | US-10-299-636-66     | Sequence 66, Appl |
| 42 | 290.5 | 62.5 | 100 | 15 | US-10-674-755-12     | Sequence 12, Appl |
| 43 | 277   | 59.6 | 195 | 15 | US-10-299-636-86     | Sequence 86, Appl |
| 44 | 273   | 58.7 | 99  | 15 | US-10-674-755-13     | Sequence 13, Appl |
| 45 | 215.5 | 46.3 | 100 | 15 | US-10-674-755-2      | Sequence 2, Appl  |

ALIGNMENTS

RESULT 1

US-10-674-755-17  
; Sequence 17, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(99)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-17

|                       |        |  |         |            |    |        |    |
|-----------------------|--------|--|---------|------------|----|--------|----|
| Query Match           | 98.3%  | Score  | 457     | DB         | 15 | Length | 99 |
| Best Local Similarity | 100.0% | Pred. No.  | 9.7e-36 |            |    |        |    |
| Matches               | 99     | Conservative   | 0       | Mismatches | 0  | Indels | 0  |
| Gaps                  | 0      |  |         |            |    |        |    |
| Qy                    | 1      | PKRINSLSQKVLKXVCRAPLQSLDAQKAEELKLEELSGKIKELDAETAELEVLQKDAE | 60      |            |    |        |    |
| Db                    | 1      | PKRINSLSQKVLKXVCRAPLQSLDAQKAEELKLEELSGKIKELDAETAELEVLQKDAE | 60      |            |    |        |    |
| Qy                    | 61     | GNNVNEATFKGLEKTTAEKKAELEAXADLKXAVDEPE                      | 99      |            |    |        |    |

D6 61 GNNVVEAYFKEGLEKTTAAEKKAELXAXADLKKAVDEPE 99

## RESULT 2

US-10-299-636-64  
; Sequence 64, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tart, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 64  
; LENGTH: 193  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae

; NAME/KEY: UNSURE  
; LOCATION: (24)  
; OTHER INFORMATION: Xaa at position 24 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (27)  
; OTHER INFORMATION: Xaa at position 27 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (98)  
; OTHER INFORMATION: Xaa at position 98 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (100)  
; OTHER INFORMATION: Xaa at position 100 is unknown  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (189)  
; OTHER INFORMATION: Xaa at position 189 is unknown  
US-10-299-636-64

Query Match 97.4%; Score 453; DB 15; Length 193;  
Best Local Similarity 99.0%; Pred. No. 5e-35;  
Matches 98; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKRTMSLSOKVXLXVCRAPIQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAE 60

D6 13 PKRTMSLSOKVXLXVCRAPIQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAE 72

Qy 61 GNNVVEAYFKEGLEKTTAAEKKAELXAXADLKKAVDEPE 99

D6 73 GNNVVEAYFKEGLEKTTAAEKKAELXAXADLKKAVDEPE 111

## RESULT 3

US-10-674-755-15  
; Sequence 15, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471

; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae

US-10-674-755-15

Query Match 80.2%; Score 373; DB 15; Length 99;  
Best Local Similarity 93.9%; Pred. No. 8.5e-28;  
Matches 77; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

D6 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

Qy 78 AEKKAELXAXADLKKAVDEPE 99

D6 78 AEKKAELXAXADLKKAVDEPE 99

## RESULT 4

US-10-674-755-14  
; Sequence 14, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; NAME/KEY: UNSURE  
; LOCATION: (1)-(99)  
; OTHER INFORMATION: amino acid 'Xaa' can be any amino acid  
US-10-674-755-14

Query Match 77.6%; Score 361; DB 15; Length 99;  
Best Local Similarity 90.2%; Pred. No. 1.2e-26;  
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

D6 18 RAPQSKLDQAQKALLKLELSGKIKELDAEIAELEVLQKDAEINNVEAYFKEGLEKTT 77

Qy 78 AEKKAELXAXADLKKAVDEPE 99

D6 78 AEKKAELXAXADLKKAVDEPE 99

## RESULT 5

US-10-299-636-69  
; Sequence 69, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan

APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 69  
LENGTH: 206  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (50)  
OTHER INFORMATION: Xaa at position 50 is unknown  
US-10-299-636-69

Query Match 77.6%; Score 361; DB 15; Length 206;  
Best Local Similarity 90.2%; Pred. No. 2.7e-26;  
Matches 74; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 18 RAPQSKLDAQKAEALLKLELSGKIKELDAEIAEVLQKDAEGNNNVEAYFKEGLEKTT 77  
DB 18 RAPQSELDTKKAKLLKLELSGKIKELDAEIAEVLQKDAEGNNNVEAYFKEGLEKTT 77  
QY 78 AEKAELEAXADLKKAVDEPE 99  
DB 78 AEKAELEKAEADLKKAVDEPE 99

RESULT 6  
US-10-299-636-75  
Sequence 75, Application US/10299636  
Publication No. US20040077847A1  
GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 75  
LENGTH: 170  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-10-299-636-75

Query Match 76.3%; Score 355; DB 15; Length 170;  
Best Local Similarity 78.8%; Pred. No. 8e-26;  
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;  
QY 2 KRIMSLSQVKLVKVCRAPLQSKLDAQKAEALLKLELSGKIKELDAEIAEVLQKDAEG 61  
DB 2 KEIDESDSELYLKEGRAPLQSKLDTKKAKLSELSKIDELDAEIAEVLQKDAEG 61  
QY 62 NNNVEAYFKEGLEKTTAEKAELEAXADLKKAVDEPE 99

Db 62 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99  
RESULT 7  
US-10-299-636-57  
Sequence 57, Application US/10299636  
Publication No. US20040077847A1  
GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 57  
LENGTH: 181  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-10-299-636-57

Query Match 76.3%; Score 355; DB 15; Length 181;  
Best Local Similarity 78.6%; Pred. No. 8.7e-26;  
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;  
QY 2 KRIMSLSQVKLVKVCRAPLQSKLDAQKAEALLKLELSGKIKELDAEIAEVLQKDAEG 61  
DB 2 KEIDESDSELYLKEGRAPLQSKLDTKKAKLSELSKIDELDAEIAEVLQKDAEG 61  
QY 62 NNNVEAYFKEGLEKTTAEKAELEAXADLKKAVDEPE 99  
DB 62 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99

RESULT 8  
US-10-299-636-95  
Sequence 95, Application US/10299636  
Publication No. US20040077847A1  
GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 95  
LENGTH: 643  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae





RESULT 13  
US-09-748-875-1  
; Sequence 1, Application US/09748875  
; Publication No. US20010016200A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/748,875  
; CURRENT FILING DATE: 2000-12-26  
; PRIOR APPLICATION NUMBER: 09/298,523  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; TYPE: PRT  
; LENGTH: 691  
; ORGANISM: Streptococcus pneumoniae  
US-09-748-875-1

Query Match 76.3%; Score 355; DB 9; Length 691;  
Best Local Similarity 78.6%; Pred. No. 4.1e-25;  
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRIMSLSQVXLYKVCRAPIQSKLDQAQKAELELLKLEELSGKIKELDAEIAELEVLQKDAEG 61  
Db 531 KEIDESDSEYDKGLRQPLQSKLDTKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 590

Qy 62 NNNVEAYFKEGLEKTTAEKKAELEAXADLKKAADLKKAVDEPE 99  
Db 591 NNNVEAYFKEGLEKTTAEKKAELEAXADLKKAADLKKAVDEPE 628

Search completed: June 21, 2005, 11:18:36  
Job time : 63.2388 secs

RESULT 14  
US-09-298-523B-1  
; Sequence 1, Application US/09298523B  
; Publication No. US20030059438A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS  
; FILE REFERENCE: 454312-3140  
; CURRENT APPLICATION NUMBER: US/09/298,523B  
; CURRENT FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; TYPE: PRT  
; LENGTH: 691  
; ORGANISM: Streptococcus pneumoniae  
US-09-298-523B-1

Query Match 76.3%; Score 355; DB 10; Length 691;  
Best Local Similarity 78.6%; Pred. No. 4.1e-25;  
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRIMSLSQVXLYKVCRAPIQSKLDQAQKAELELLKLEELSGKIKELDAEIAELEVLQKDAEG 61  
Db 531 KEIDESDSEYDKGLRQPLQSKLDTKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 590

Qy 62 NNNVEAYFKEGLEKTTAEKKAELEAXADLKKAADLKKAVDEPE 99  
Db 591 NNNVEAYFKEGLEKTTAEKKAELEAXADLKKAADLKKAVDEPE 628

RESULT 15  
US-09-748-875-62  
; Sequence 62, Application US/09748875  
; Publication No. US20010016200A1  
; GENERAL INFORMATION:  
; APPLICANT: BRILES et al.

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 9.9 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PIRMSLSQVXKXVCRAP.....KKAEXAXADLKXAVDEPE 99

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 304   | 65.4        | 619    | 2 A97887 | surface protein ps |
| 2          | 304   | 65.4        | 619    | 2 A41971 | surface protein ps |
| 3          | 110.5 | 23.8        | 744    | 2 P95013 | pneumococcal surfa |
| 4          | 103   | 22.2        | 161    | 2 S48396 | tropomyosin TPM2 - |
| 5          | 103   | 22.2        | 281    | 2 F75216 | hypothetical prote |
| 6          | 101.5 | 21.8        | 1006   | 2 C70445 | ATPase subunit of  |
| 7          | 100   | 21.5        | 1312   | 2 T30845 | probable DNA repai |
| 8          | 95    | 20.4        | 785    | 2 T01025 | hypothetical prote |
| 9          | 94    | 20.2        | 229    | 2 S70532 | outer surface prot |
| 10         | 94    | 20.2        | 1808   | 2 T15099 | hypothetical prote |
| 11         | 93    | 20.0        | 876    | 2 A23767 | myosin heavy chain |
| 12         | 92    | 19.8        | 395    | 2 AC1754 | capsid protein [ba |
| 13         | 92    | 19.8        | 1171   | 2 T45706 | chromosome-associ  |
| 14         | 91.5  | 19.7        | 852    | 2 D72230 | conserved hypotet  |
| 15         | 91.5  | 19.7        | 1269   | 2 P84730 | probable myosin he |
| 16         | 91    | 19.6        | 880    | 2 F75103 | conserved hypotet  |
| 17         | 90.5  | 19.5        | 1976   | 2 A59252 | myosin heavy chain |
| 18         | 90    | 19.4        | 1319   | 2 A28313 | skel protein - fr  |
| 19         | 90    | 19.4        | 1938   | 2 A59293 | skeletal myosin he |
| 20         | 89.5  | 19.2        | 339    | 2 E71169 | hypothetical prote |
| 21         | 89    | 19.1        | 784    | 2 T05409 | hypothetical prote |
| 22         | 88.5  | 19.0        | 2116   | 2 A26655 | myosin heavy chain |
| 23         | 88.5  | 19.0        | 3187   | 2 JC5837 | 364K Golgi complex |
| 24         | 88    | 18.9        | 955    | 2 S24348 | myosin heavy chain |
| 25         | 88    | 18.9        | 1938   | 1 JX0178 | myosin heavy chain |
| 26         | 88    | 18.9        | 1940   | 2 A29320 | myosin heavy chain |
| 27         | 88    | 18.9        | 1974   | 2 T30010 | hypothetical prote |
| 28         | 88    | 18.9        | 4574   | 2 G02520 | plectin - human    |
| 29         | 87.5  | 18.8        | 279    | 2 D71453 | hypothetical prote |

RESULT 1

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C:Accession: A97887

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: A97887

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <XUR>

A:Cross-References: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl

C:Genetics:

A:Gene: pspA

Query Match 65.4%; Score 304; DB 2; Length 619;  
Best Local Similarity 80.5%; Pred. No. 3.8e-16;  
Matches 66; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 18 RAPLOSKLDAQKAEILLKLEELSGKIKELDAETAELEVLKDAEGNNVAYPKGLEKTT 77

Db 240 RAPLOSKLDAQKAEILLKLEELSGKIKELDAETAELEVLKDAEGNNVAYPKGLEKTTI 299

QY 78 AEKKAEXAXADLKXAVDEPE 99

Db 300 AAKKAEXLEKTEADLKXAVNEPE 321

RESULT 2

A41971

surface protein pspA precursor - Streptococcus pneumoniae

N:Alternate names: pneumococcal surface protein A

C:Species: Streptococcus pneumoniae

C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A41971; A60282; A33134

R:Yother, J.; Briles, D.E.

J. Bacteriol. 174, 601-609, 1992

A>Title: Structural properties and evolutionary relationships of PspA, a surface protein

A:Reference number: A41971; MUID:92105030; PMID:1729249

A:Accession: A41971

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-619 <YOT>

A:Cross-References: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:G153840; PIDN:AAA2701f

A>Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIIP:75636)

R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.

N;Alternate names: protein YII138c  
C;Species: Saccharomyces cerevisiae  
C;Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 09-Jul-2004  
C;Accession: S48396; A56490  
R;Churcher, C.  
submitted to the EMBL Data Library, September 1994  
A;Reference number: S48310  
A;Accession: S48396  
A;Molecule type: cDNA  
A;Residues: 1-161 <CHU>  
A;Cross-references: UNIPROT:P40414; GB:Z47047; EMBL:Z38059; NID:g603997; PID:g763208; MIPS:  
R;Drees, B.; Brown, C.; Barrrell, B.G.; Bretscher, A.  
J. Cell Biol. 128, 383-392, 1995  
A;Title: Tropomyosin is essential in yeast, yet the TPM1 and TPM2 products perform distribut  
A;Reference number: A56490; MUID:95146545; PMID:7844152  
A;Accession: A56490  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-161 <DRE>  
A;Cross-references: GB:Z47047; GB:Z38059; NID:g603997; PID:g763208  
C;Genetics:  
A;Gene: SGD:TPM2  
A;Cross-references: SGD:S0001400; MIPS:YII138C  
A;Map position: 9L  
C;Superfamily: tropomyosin TPM1  
C;Keywords: cytoskeleton

Query Match 22.2%; Score 103; DB 2; Length 161;  
Best Local Similarity 31.9%; Pred. No. 0.26; Matches 29; Conservative 20; Mismatches 26; Indels 16; Gaps 3;

Qy 22 QSKLDAOKAEILLKEE-----LSGKIKELDAETAELEVLQKD---AEGNNNVE 66  
Db 19 QEKYEELEQLKELQSQNTKENETKSISAKNEQLDSVEVKLESQSLDTKQLAEDSNLR 78

Qy 67 AYFKEGLEKTTAETKAELXAXADLKKAUDE 97  
Db 79 SN-NENTYTKNQDLQEQLDSEAKLKEAMDK 108

RESULT 5  
F75216  
hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: F75216  
R;anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc  
A;Reference number: A75001  
A;Accession: F75216  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-281 <KAW>  
A;Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49418J  
C;Genetics:  
A;Gene: PAB2181

Query Match 22.2%; Score 103; DB 2; Length 281;  
Best Local Similarity 30.5%; Pred. No. 0.45; Matches 29; Conservative 25; Mismatches 23; Indels 18; Gaps 3;

Qy 21 LQSKLDAQKAEIL-----KLSELSKIELDAETAELEVLQKDAGN-----NN 64  
Db 146 VKAKIEAAELSELKAKKIEELKGKVLEQEKKEKKLESEVKLMYEAKRAEE 205

Qy 65 VEAYFKEGLEKTTAETKAELXAXADLKKAUDE 99  
Db 206 LEAKLYEEKVKREE--ELERKVSLELRSLNEYE 238

RESULT 6

RESULT 10  
T15099  
hypothetical protein W03F8.5 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 20-Sep-1999 #sequence revision 20-Sep-1999 #text change 09-Jul-2004



R.Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing  
A:Reference number: A72200; PMID:99287316; PMID:10360571

A:Accession: D72230  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-852 <ARN>  
A:Cross-references: UNIPROT:Q9X1X1; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3670  
A:Experimental source: strain MSB8

C:Genetics:  
A:Gene: Tm1636

C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 19.7%; Score 91.5; DB 2; Length 852;  
Best Local Similarity 31.1%; Pred. No. 10;  
Matches 23; Conservative 18; Mismatches 22; Indels 11; Gaps 2;

QY 21 LOSKLDQAQKAEELKLEE---LSGRIKIKELDAETAELEVLQKDAEGNNVVEAYFKEGLEKT 76

DB 521 LEEKLDEKRLKRIEERHSISQKITAADVQISQIENQLKEIKGE-----IEAKRET 573

QY 77 TAEKKAELXAXAD 90

DB 574 LKEQREMDQLKSD 587

#### RESULT 15

F84730

probable myosin heavy chain [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: F84730

R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; PMID:20083487; PMID:10617197

A:Accession: F84730

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1269 <STO>

A:Cross-references: GB:AE002093; NID:g6598483; PIDN:AAC69932.2; GSPDB:GN00139

C:Genetics:

A:Gene: At2g32240

A:Map position: 2

Query Match 19.7%; Score 91.5; DB 2; Length 1269;  
Best Local Similarity 27.0%; Pred. No. 15;  
Matches 27; Conservative 22; Mismatches 48; Indels 3; Gaps 1;

QY 3 RIMSLQKVKLVKVCRAPIQSKLDQAQKAEELKLEEISGRIKELDAETAELEVLQKDAEGN 62

DB 300 RLLETRQKVSSTEALIDELTQLEQKKAESRPFKEELSVLQDLDAQTGKLQAKLSQEGL 359

QY 63 NN---VEAYFKEGLEKTAEKKAELXAXADLKKAVDEPE 99

DB 360 NSKLAEELKEELLESLSKQDEKRLTANEKLAELVLEKE 399

Search completed: June 21, 2005, 10:12:03

Job time : 10.9 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time: 61.3194 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-17

Perfect score: 465

Sequence: 1 PKRMSLSQVKKXVCRAP.....KKAELEXAXADLKXAVDEPE 99

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID    | Description        |
|------------|-------|---------------|--------|----------|--------------------|
| 1          | 373   | 80.2          | 395    | 2 Q9LAY2 | Q9lay2 streptococc |
| 2          | 373   | 80.2          | 408    | 2 Q9LAY0 | Q9lay0 streptococc |
| 3          | 367   | 78.9          | 249    | 2 Q9L575 | Q9l575 streptococc |
| 4          | 366   | 78.7          | 426    | 2 Q9LAY5 | Q9lay5 streptococc |
| 5          | 365   | 78.5          | 224    | 2 Q8GNS8 | Q8gns8 streptococc |
| 6          | 362   | 77.8          | 99     | 2 Q8KQK4 | Q8kqk4 streptococc |
| 7          | 355   | 76.3          | 739    | 2 Q9RQT4 | Q9rqt4 streptococc |
| 8          | 355   | 76.3          | 820    | 2 Q9RQT1 | Q9rqt1 streptococc |
| 9          | 355   | 76.3          | 929    | 2 Q9KK19 | Q9kk19 streptococc |
| 10         | 355   | 76.3          | 929    | 2 Q9ZAY5 | Q9zay5 streptococc |
| 11         | 336   | 72.3          | 437    | 2 Q9LAY4 | Q9lay4 streptococc |
| 12         | 315.5 | 67.8          | 869    | 2 Q9KK27 | Q9kk27 streptococc |
| 13         | 304   | 65.4          | 619    | 2 Q54972 | Q54972 streptococc |
| 14         | 304   | 65.4          | 619    | 2 Q8DR10 | Q8dr10 streptococc |
| 15         | 297   | 63.9          | 417    | 2 Q9LAY3 | Q9lay3 streptococc |
| 16         | 282   | 60.6          | 415    | 2 Q9LAY1 | Q9lay1 streptococc |
| 17         | 221.5 | 47.6          | 246    | 2 Q9L578 | Q9l578 streptococc |
| 18         | 219.5 | 47.2          | 194    | 2 Q9L5B5 | Q9l5b5 streptococc |
| 19         | 219.5 | 47.2          | 218    | 2 Q6UEB2 | Q6ueb2 streptococc |
| 20         | 219.5 | 47.2          | 233    | 2 Q9L568 | Q9l568 streptococc |
| 21         | 219.5 | 47.2          | 236    | 2 Q9L569 | Q9l569 streptococc |
| 22         | 219.5 | 47.2          | 243    | 2 Q9L564 | Q9l564 streptococc |
| 23         | 219.5 | 47.2          | 243    | 2 Q9L567 | Q9l567 streptococc |
| 24         | 219.5 | 47.2          | 244    | 2 Q9L565 | Q9l565 streptococc |
| 25         | 219.5 | 47.2          | 247    | 2 Q9L566 | Q9l566 streptococc |
| 26         | 219.5 | 47.2          | 249    | 2 Q9L570 | Q9l570 streptococc |
| 27         | 219.5 | 47.2          | 254    | 2 Q9L563 | Q9l563 streptococc |
| 28         | 219.5 | 47.2          | 401    | 2 Q9LAZ2 | Q9laz2 streptococc |
| 29         | 217.5 | 46.8          | 222    | 2 Q9L577 | Q9l577 streptococc |
| 30         | 217.5 | 46.8          | 262    | 2 Q9L576 | Q9l576 streptococc |
| 31         | 217.5 | 46.8          | 415    | 2 Q9LAY7 | Q9lay7 streptococc |

```
32 216.5 46.6 225 2 Q9L591 streptococc
33 215.5 46.3 394 2 Q9LAY6 streptococc
34 215.5 46.3 395 2 Q9LAZ1 streptococc
35 214.5 46.1 393 2 Q9LAZ3 streptococc
36 210.5 45.3 416 2 Q9LAY8 streptococc
37 209.5 45.1 255 2 Q9L581 streptococc
38 209.5 45.1 255 2 Q9L5B6 streptococc
39 206.5 44.4 406 2 Q9LAZ0 streptococc
40 203.5 43.8 237 2 Q9L592 streptococc
41 203.5 43.8 395 2 Q9LAY9 streptococc
42 198.5 42.7 340 2 Q8KQK5 streptococc
43 194.5 41.8 207 2 Q8GNS9 streptococc
44 160 34.4 107 2 Q8KQK2 streptococc
45 160 34.4 246 2 Q9L5B4 streptococc
```

#### ALIGNMENTS

RESULT 1

```
Q9LAY2
ID Q9LAY2 PRELIMINARY; PRT; 395 AA.
AC Q9LAY2;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DE PspA (Fragment).
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF6796;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071813; AAF27709.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR011047; Quin_abc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYSIN.
FT NON_TER 395
SQ SEQUENCE 395 AA; 42963 MW; 58B6EF956BCBCC1E CRC64;

Query Match 80.2%; Score 373; DB 2; Length 395;
Best Local Similarity 93.9%; Pred. No. 1.4e-19;
Matches 77; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 18 RAPLQSLDAQKAEILKLELSGKIKELDAEIALEVLQKDAEGNNVAYFKEGLEKTT 77
Db 242 RAPLQSLDAKAKLKLKLELSGKIEELDAEIALEVLQKDAEGNNVAYFKEGLEKTT 301
QY 78 AEKKAEXAXADLKXAVDEPE 99
Db 302 AEKKAEXAXADLKXAVDEPE 323
```

RESULT 2

```
Q9LAY0
ID Q9LAY0 PRELIMINARY; PRT; 408 AA.
AC Q9LAY0;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DE PspA (Fragment).
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
```

```

OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071815; AAF27711.1; -.
DR InterPro; IPR009053; Pfefoldin.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 408 408
SQ SEQUENCE 408 AA; 44254 MW; 4F64D874217297EF CRC64;

Query Match 80.2%; Score 373; DB 2; Length 408;
Best Local Similarity 93.9%; Pred. No. 1.4e-19;
Matches 77; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 18 RAPQSKLDQAQKAELELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKGLEKTT 77
Db 245 RAPQSKLDQAQKAELELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKGLEKTT 304

Qy 78 AEKKAELXAXADLKKAVDPE 99
Db 305 AEKKAELXAXADLKKAVDPE 326

RESULT 3
Q9L575 PRELIMINARY; PRT; 249 AA.
AC Q9L575;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
pneumococcal strains in the United States and of internationally
disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255552; AAF68105.1; -.
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 26986 MW; 7916D5014E387BD8 CRC64;

Query Match 78.9%; Score 367; DB 2; Length 249;
Best Local Similarity 79.6%; Pred. No. 2.4e-19;
Matches 78; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 2 KRIMSLSQKVLKXVCRAPQSKLDQAQKAELELSGKIKELDAIEAEVQLKDAEG 61
Db 75 KEIDSDSDYIKGFRAPQSKLDQAQKAELELSGKIKELDAIEAEVQLKDAEG 134

Qy 62 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDPE 99
Db 135 NNNVEAYFKGLEKTTAEKKAELXAXADLKKAVDPE 172

us-10-674-755-17.rup

RESULT 4
Q9LAYS PRELIMINARY; PRT; 426 AA.
AC Q9LAYS;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBL5;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071810; AAF27706.1; -.
DR HSP; P00192; IM6T.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 426 426
SQ SEQUENCE 426 AA; 46534 MW; 81AA11348CBE6634 CRC64;

Query Match 78.7%; Score 366; DB 2; Length 426;
Best Local Similarity 87.4%; Pred. No. 4.8e-19;
Matches 76; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 13 LKXVCRAPQSKLDQAQKAELELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKEG 72
Db 227 VKESLRAPQSKLDQAQKAELELSGKIKELDAIEAEVQLKDAEGNNVVEAYFKEG 286

Qy 73 LEKTTAEKKAELXAXADLKKAVDPE 99
Db 287 LEKTTAEKKAELXAXADLKKAVDPE 313

RESULT 5
Q8GNS8 PRELIMINARY; PRT; 224 AA.
AC Q8GNS8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PN124;
RX MEDLINE=22241996; PubMed=12354862;
RA Dicuonzo G., Gherardi G., Gertz R.E., D'Ambrosio F., Goglio A.,
RA Lorino G., Recchia S., Pantosti A., Beall B.;
RT "genotypes of invasive pneumococcal isolates recently recovered from
Italian patients.";
RL J. Clin. Microbiol. 40:3660-3665(2002).
DR EMBL; AF490267; AAN37735.1; -.
DR HSP; P00192; JAPC.
DR InterPro; IPR009082; His_kin_homodim.
FT NON_TER 224 224
SQ SEQUENCE 224 AA; 23418 MW; 48674E27AFB66A95 CRC64;

```

```
Query Match 78.5%; Score 365; DB 2; Length 224;
Best Local Similarity 91.5%; Pred. No. 3.1e-19;
Matches 75; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 18 RAPQSLQDAQKAEELKLELSKIKELDAEIAEVLQKDAEGNNVAYFKEGLEKTT 77
DB 34 RAPQSELDTKKAKLKLKLELSKIKELDAEIAEVLQKDAEGNNVAYFKEGLEKTT 93
QY 78 AEKKAEXAXADLKXAVDEPE 99
DB 94 AEKKAELKAEADLKXAVDEPE 115

RESULT 6
Q8KQK4 PRELIMINARY; PRT; 99 AA.
AC Q8KQK4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=371/00;
RX MEDLINE=22170754; PubMed=12183557;
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082388; AAL92493.1; -.
FT NON TER 1
FT NON TER 99
FT NON TER 99
SQ SEQUENCE 99 AA; 11105 MW; 7A13308CA174A3A7 CRC64;

Query Match 77.8%; Score 362; DB 2; Length 99;
Best Local Similarity 78.6%; Pred. No. 2.3e-19;
Matches 77; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 2 KRIMSLSQVKLXKVCRAPLQSKLDAQKAEELKLELSKIKELDAEIAEVLQKDAEG 61
DB 2 KEIDESDSEYKGLRAPLQSLDTKKAKLLKLELSGKIEELDAEIAEVLQKDAEG 61
QY 62 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKXAVDEPE 99
DB 62 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKXAVDEPE 99

RESULT 7
Q9RQT4 PRELIMINARY; PRT; 739 AA.
AC Q9RQT4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EI34;
RX MEDLINE=20038319; PubMed=10569772;

Query Match 76.3%; Score 355; DB 2; Length 820;
Best Local Similarity 78.6%; Pred. No. 5.6e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;
```

```
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068647; AAF13457.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON TER 739
FT NON TER 739
SQ SEQUENCE 739 AA; 83960 MW; 7EB2F2F676ABF989 CRC64;

Query Match 76.3%; Score 355; DB 2; Length 739;
Best Local Similarity 78.6%; Pred. No. 5.1e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 KRIMSLSQVKLXKVCRAPLQSKLDAQKAEELKLELSKIKELDAEIAEVLQKDAEG 61
DB 538 KEIDESDSEYKGLRAPLQSLDTKKAKLLKLELSGKIEELDAEIAEVLQKDAEG 597
QY 62 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKXAVDEPE 99
DB 598 NNNVEAYFKEGLEKTTAEKKAELKAELEXAXADLKXAVDEPE 635

RESULT 8
Q9RQT1 PRELIMINARY; PRT; 820 AA.
AC Q9RQT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG9163;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068650; AAF13460.1; -.
DR HSSP; P04268; IIC2.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON TER 820
FT NON TER 820
SQ SEQUENCE 820 AA; 91752 MW; 33C095849ABB0942 CRC64;

Query Match 76.3%; Score 355; DB 2; Length 820;
Best Local Similarity 78.6%; Pred. No. 5.6e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;
```

```

Qy 2 KRMSLSQKVKLKKVCRAPIQSKLDAQKAEALLKLEELSGKIKELDAEIAELEVLQKDAEG 61
Db 531 KEIDESDSEDYKLEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 590
Qy 62 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
Db 591 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 9
Q9KK19 ID Q9KK19 PRELIMINARY; PRT; 929 AA.
AC Q9KK19
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=8710;
RX MEDLINE=2188621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
ST Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154037; AAF73809.1; -.
DR HSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR Pfam; PF01473; CW_binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 105003 MW; 2DC8293302FAFA64 CRC64;

Query Match 76.3%; Score 355; DB 2; Length 929;
Best Local Similarity 78.6%; Pred. No. 6.4e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRMSLSQKVKLKKVCRAPIQSKLDAQKAEALLKLEELSGKIKELDAEIAELEVLQKDAEG 61
Db 531 KEIDESDSEDYKLEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 590
Qy 62 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
Db 591 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 10
Q9ZAY5 ID Q9ZAY5 PRELIMINARY; PRT; 929 AA.
AC Q9ZAY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein C.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=EF6796;
RX MEDLINE=20038319; PubMed=10569772;

```

```

RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, pspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; U72655; AAD00184.1; -.
DR HSP; P06653; 1HXC.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 11.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 929 AA; 104991 MW; 2DC8293302FFB081 CRC64;

Query Match 76.3%; Score 355; DB 2; Length 929;
Best Local Similarity 78.6%; Pred. No. 6.4e-18;
Matches 77; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 KRMSLSQKVKLKKVCRAPIQSKLDAQKAEALLKLEELSGKIKELDAEIAELEVLQKDAEG 61
Db 531 KEIDESDSEDYKLEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAELEVLQKDAEG 590
Qy 62 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 99
Db 591 NNNVEAYFKEGLEKTTAEKAELEKAEADLKKAVDEPE 628

RESULT 11
Q9LAY4 ID Q9LAY4 PRELIMINARY; PRT; 437 AA.
AC Q9LAY4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=E134;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071811; AAF27707.1; -.
DR InterPro; IPR002479; CW_binding.
DR Pfam; PF01473; CW_binding_1; 1.
DR NON TER 437
SQ SEQUENCE 437 AA; 48071 MW; F6EFD2CD13E08CD8 CRC64;

Query Match 72.3%; Score 336; DB 2; Length 437;
Best Local Similarity 86.6%; Pred. No. 7.6e-17;
Matches 71; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 18 RAPLQSKLDAQKAEALLKLEELSGKIKELDAEIAELEVLQKDAEGNNNVEAYFKEGLEKTT 77
Db 252 RAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAEKHVVVQLKDAEGNNNVEAYFKEGLEKTT 311
Qy 78 AEKAELEKAEADLKKAVDEPE 99
Db 312 AEKAELEKAEADLKKAVDEPE 333

RESULT 12

```

```
Q9KK27
ID Q9KK27 PRELIMINARY; PRT; 869 AA.
AC Q9KK27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=95;
RC MEDLINE=21188621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Tannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154032; AAF73801.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW binding 1; 8.
DR Pfam; PF05062; RICH; 2.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk signal; 1.
SQ SEQUENCE 869 AA; 98732 MW; AFF2B504347E0220 CRC64;

Query Match 67.4%; Score 315.5; DB 2; Length 869;
Best Local Similarity 72.4%; Pred. No. 4.6e-15;
Matches 71; Conservative 3; Mismatches 23; Indels 1; Gaps 1;

QY 2 KRIMLSQKXKVCRAPOSKLDAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEG 61
DB 538 KEIDESDSDEYKLEGURAPQSKLDTKKAKLSKLEELSDKIDELVY-CNLRSQLKDAEG 596

QY 62 NNNVEAYFKEGLEKTTAEKKAELKXAXADLKKAVIDEPE 99
DB 597 NNNVEAYFKEGLEKTTAEKKAELKXAXADLKKAVIDEPE 634

RESULT 13
Q54972
ID Q54972 PRELIMINARY; PRT; 619 AA.
AC Q54972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A precursor.
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_SEQUENCE FROM N.A.
RP MEDLINE=92105030; PubMed=1729249;
RX Yother J., Briles D.E.;
RT "Structural properties and evolutionary relationships of PspA, a
RT surface protein of Streptococcus pneumoniae, as revealed by sequence
RT analysis.";
RL J. Bacteriol. 174:601-609(1992).
RN [2]_SEQUENCE FROM N.A.
RP Yother J., Briles D.E.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M74122; AAA27018.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
```

```
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW SIGNAL.
FT SIGNAL. 1 31 Potential.
FT CHAIN 32 619 pneumococcal surface protein A.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 65.4%; Score 304; DB 2; Length 619;
Best Local Similarity 80.5%; Pred. No. 2.3e-14;
Matches 66; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 18 RAPLOSKLDAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTT 77
DB 240 RAPLOSKLDAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTTI 299

QY 78 AEKKAELKXAXADLKKAVIDEPE 99
DB 300 AAKKAELKTEADLKKAVIDEPE 321

RESULT 14
Q8DR10
ID Q8DR10 PRELIMINARY; PRT; 619 AA.
AC Q8DR10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Surface protein pspA.
GN Name=pspA; OrderedLocNames=sp0121;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]_SEQUENCE FROM N.A.
RP MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Barget S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.N., McHenry M., McLeaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008396; AAK98925.1; -.
DR PIR; A41971; A41971.
DR PIR; A97887; A97887.
DR HSSP; P06653; 1HCX.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_3.
KW Complete proteome.
SQ SEQUENCE 619 AA; 68605 MW; 5AA8BDB40C2841CA CRC64;

Query Match 65.4%; Score 304; DB 2; Length 619;
Best Local Similarity 80.5%; Pred. No. 2.3e-14;
Matches 66; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 18 RAPLOSKLDAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTT 77
DB 240 RAPLOSKLDAQKAEKLLKLEELSGKIKELDAEIAEVLQKDAEGNNVYAYFKEGLEKTTI 299

QY 78 AEKKAELKXAXADLKKAVIDEPE 99
DB 300 AAKKAELKTEADLKKAVIDEPE 321
```

```
RESULT 15
Q9LAY3 PRELIMINARY; PRT; 417 AA.
AC Q9LAY3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF10197;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K.; Becker R.; Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071812; AAF27708.1; -.
DR HSP; P00192; 256B.
FT NON TER 417
SQ SEQUENCE 417 AA; 46960 MW; 876EAD3276506EEC CRC64;

Query Match 63.9%; Score 297; DB 2; Length 417;
Best Local Similarity 68.4%; Pred. No. 5.le-14;
Matches 67; Conservative 8; Mismatches 23; Indels 0; Gaps 0;

Qy 2 KRIMSLSQVXLKXVCRAPLQSKLDAQKAEILKLEELSGKIKELDAEIAELEVLQKDAEG 61
Db 214 KEIDSESDYVXGFGFRAPLQSELDQAQKLSLEELSDKIDELDAEIAKLEDLQKAAEE 273

Qy 62 NNNVEAYFKEGLEKTTAEKKAEXAXADLKKAVIDEPE 99
Db 274 NNNVEDYFKEGLEKTTIAKKALEKTEADLKKAVIDEPE 311
```

Search completed: June 21, 2005, 10:22:13  
Job time : 62.3194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 76.0837 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-18

Perfect score: 505

Sequence: 1 LAKQTELEKLLDPEGKT.....TQKELDALNELGPDGDEE 102

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 469   | 92.9        | 213    | 7     | ABW02601 Bg8090c p  |
| 2          | 469   | 92.9        | 8991   | 6     | ABU08487 S. pneumo  |
| 3          | 468   | 92.7        | 416    | 8     | ADK52498 alpha hel  |
| 4          | 468   | 92.7        | 526    | 8     | ADK52497 PspA mole  |
| 5          | 468   | 92.7        | 744    | 6     | ABU00449 S. pneumo  |
| 6          | 468   | 92.7        | 744    | 8     | ADM52054 S. pneumo  |
| 7          | 468   | 92.7        | 745    | 3     | AAy81652 Streptoco  |
| 8          | 467   | 92.5        | 213    | 2     | AAW14567 Streptoco  |
| 9          | 464   | 91.9        | 641    | 5     | AAW61217 Streptoco  |
| 10         | 464   | 91.9        | 641    | 5     | ABP54636 S. pneumo  |
| 11         | 464   | 91.9        | 641    | 7     | ADC45241 S. pneumo  |
| 12         | 463   | 91.7        | 197    | 7     | ABW02598 Ac122c pn  |
| 13         | 442   | 87.5        | 233    | 7     | ABW02606 Bf1019c pn |
| 14         | 439.5 | 87.0        | 196    | 2     | AAW14564 Streptoco  |
| 15         | 435   | 86.1        | 233    | 2     | AAW14572 Streptoco  |
| 16         | 298   | 59.0        | 211    | 7     | ABW02621 Bg11703c   |
| 17         | 298   | 59.0        | 238    | 2     | AAW14587 Streptoco  |
| 18         | 293   | 58.0        | 232    | 7     | ABW02624 Ef5668c p  |
| 19         | 293   | 58.0        | 275    | 8     | ADOS2055 S. pneumo  |
| 20         | 293   | 58.0        | 369    | 8     | ADK52496 alpha hel  |
| 21         | 293   | 58.0        | 458    | 2     | AAW14592 Streptoco  |
| 22         | 293   | 58.0        | 458    | 7     | ABW02626 Ef5668 pn  |
| 23         | 293   | 58.0        | 653    | 8     | ADK52495 PspA mole  |
| 24         | 293   | 58.0        | 653    | 8     | ADOS2080 S. pneumo  |
| 25         | 283   | 56.0        | 212    | 2     | AAW14588 Streptoco  |

|    |       |      |     |   |          |                     |
|----|-------|------|-----|---|----------|---------------------|
| 26 | 283   | 56.0 | 212 | 7 | ABW02622 | Abw02622 Bg7817c p  |
| 27 | 282.5 | 55.9 | 233 | 2 | AAW14590 | Aaw14590 Streptoco  |
| 28 | 253   | 50.1 | 459 | 8 | ADO15316 | Ado15316 S. pneumon |
| 29 | 240.5 | 47.6 | 184 | 2 | AAW14589 | Aaw14589 Streptoco  |
| 30 | 236   | 46.7 | 185 | 7 | ABW02623 | Abw02623 Bg7561c p  |
| 31 | 178   | 35.2 | 487 | 8 | ADR04321 | Adr04321 Streptoco  |
| 32 | 178   | 35.2 | 489 | 8 | ADO52088 | Ado52088 Streptoco  |
| 33 | 178   | 35.2 | 524 | 8 | ADO52082 | Ado52082 E. coli B  |
| 34 | 178   | 35.2 | 627 | 8 | ADO52129 | Ado52129 E. coli B  |
| 35 | 177.5 | 35.1 | 119 | 2 | AAW46291 | Aaw46291 Pneumonoc  |
| 36 | 177.5 | 35.1 | 215 | 2 | AAW14563 | Aaw14563 Streptoco  |
| 37 | 177.5 | 35.1 | 215 | 7 | ABW02597 | Abw02597 Atcc6303c  |
| 38 | 165.5 | 32.8 | 290 | 8 | ADO52119 | Ado52119 PVA3637 b  |
| 39 | 165.5 | 32.8 | 298 | 8 | ADO52127 | Ado52127 PFA3637 b  |
| 40 | 160   | 31.7 | 230 | 8 | ADO52086 | Ado52086 S. pneumo  |
| 41 | 160   | 31.7 | 230 | 8 | ADR04319 | Adr04319 Streptoco  |
| 42 | 126   | 25.0 | 550 | 8 | ADK48356 | Adk48356 Streptoco  |
| 43 | 126   | 25.0 | 550 | 8 | ADR95223 | Adr95223 Novel S.   |
| 44 | 126   | 25.0 | 623 | 6 | ABU08494 | Abu08494 Fragment   |
| 45 | 121.5 | 24.1 | 204 | 2 | AAW14571 | Aaw14571 Streptoco  |

#### ALIGNMENTS

##### RESULT 1

ABW02601

ID ABW02601 standard; protein; 213 AA.

XX AC ABW02601;

XX DT 12-FEB-2004 (first entry)

XX DE Bg8090c pneumococcal surface protein A (PspA) central region.

XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

XX KW immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT Misc-difference 2 /label= Unknown

XX PN US6592876-B1.

XX PD 15-JUL-2003.

XX PF 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX PA (UABR-) UAB RES FOUND.

XX PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX DR WPI; 2003-862841/80.

XX PT Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

XX PS Example 6; SEQ ID NO 47; 121pp; English.

XX CC The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspA) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, CC

CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Bg8090c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 SQ Sequence 213 AA;  
 Query Match 92.9%; Score 469; DB 7; Length 213;  
 Best Local Similarity 95.2%; Pred. No. 2.4e-35;  
 Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
 QY 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 58  
 |||||  
 DB 59 LAKKQTELEKLLDNDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 118  
 |||||  
 QY 59 GADSEDDTAALPNKLTATKAELEKTQKELDAALNELGPDGDEEE 102  
 |||||  
 DB 119 GADPEDDTAALPNKLTATKAEFEKTPKELDAALNELGPDGDEEE 162  
 |||||  
 RESULT 2  
 ABU08487  
 ID ABU08487 standard; protein; 8991 AA.  
 AC  
 AC ABU08487;  
 DT 24-JUN-2003 (first entry)  
 XX  
 XX S. pneumoniae pneumococcal surface protein A (PspA) protein.  
 XX  
 XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.  
 XX  
 XX Streptococcus pneumoniae.  
 XX  
 PH Key Location/Qualifiers  
 FT Misc-difference 1..8991  
 FT /note= "All Xaa residues within this sequence are  
 FT unknown"  
 XX  
 XX US6500613-B1.  
 XX  
 XX 31-DEC-2002.  
 XX  
 XX 16-SEP-1996; 96US-00714741.  
 XX  
 XX 15-SEP-1995; 95US-00529055.  
 XX  
 XX (UYAL-) UNIV ALABAMA.  
 XX  
 XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 XX WPI; 2003-361534/34.  
 XX  
 XX Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.  
 XX  
 XX Disclosure; Col 145-188; 186pp; English.  
 XX  
 XX The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents S. pneumoniae

CC PspA protein  
 XX  
 SQ Sequence 8991 AA;  
 Query Match 92.9%; Score 469; DB 6; Length 8991;  
 Best Local Similarity 95.2%; Pred. No. 2e-33;  
 Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
 QY 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 58  
 |||||  
 DB 4625 LAKKQTELEKLLDNDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 4684  
 |||||  
 QY 59 GADSEDDTAALPNKLTATKAELEKTQKELDAALNELGPDGDEEE 102  
 |||||  
 DB 4685 GADPEDDTAALPNKLTATKAEFEKTPKELDAALNELGPDGDEEE 4728  
 |||||  
 RESULT 3  
 ADK52498  
 ID ADK52498 standard; protein; 416 AA.  
 XX  
 AC ADK52498;  
 XX  
 XX 20-MAY-2004 (first entry)  
 XX  
 XX alpha helical region PspA molecule from the EF3296 strain.  
 DE  
 XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;  
 KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;  
 KW Hodgkin's disease.  
 XX  
 XX Streptococcus pneumoniae.  
 OS  
 XX WO2004016231-A2.  
 PN  
 XX 26-FEB-2004.  
 PD  
 XX 17-FEB-2003; 2003WO-US008199.  
 PF  
 XX 15-MAR-2002; 2002US-0365351P.  
 PR  
 XX (UABR-) UAB RES FOUND.  
 PA  
 XX Briles DE;  
 PI  
 XX WPI; 2004-192068/18.  
 DR  
 XX Treating Streptococcus pneumoniae infection in a subject lacking a  
 PT functional spleen comprises administering an antibody that recognizes  
 PT pneumococcal surface protein A (PspA) or its binding portion.  
 XX  
 XX Claim 17; SEQ ID NO 4; 41pp; English.  
 XX  
 XX The present invention relates to treating Streptococcus pneumoniae  
 CC infection in a subject lacking a functional spleen comprises  
 CC administering an antibody that recognizes pneumococcal surface protein A  
 CC (PspA) or its binding portion. The method is useful for treating or  
 CC preventing Streptococcus pneumoniae infection in a subject lacking a  
 CC functional spleen. The disease-associated injury is especially due to  
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell  
 CC anemia or Hodgkin's disease. The present sequence represents the alpha  
 CC helical region PspA molecule from the EF3296 strain of Streptococcus  
 CC pneumoniae.  
 XX  
 XX Sequence 416 AA;  
 SQ  
 Query Match 92.7%; Score 468; DB 8; Length 416;  
 Best Local Similarity 96.2%; Pred. No. 6.5e-35;  
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
 QY 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 58  
 |||||  
 DB 241 LAKKQTELEKLLDNDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 300  
 |||||



QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEE 102  
 |||||  
 Db 301 GADSEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEE 344

RESULT 4  
 ADK52497  
 ID ADK52497 standard; protein; 526 AA.

XX AC ADK52497;  
 XX 20-MAY-2004 (first entry)  
 XX PepA molecule from the EF3296 strain of Streptococcus pneumoniae.

DE Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;  
 XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;  
 KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;  
 KW Hodgkin's disease.

XX Streptococcus pneumoniae.

OS WO2004016231-A2.

PN 26-FEB-2004.

PD 17-FEB-2003; 2003WO-US008199.

PF 15-MAR-2002; 2002US-0365351P.

PR (UABR-) UAB RES FOUND.

XX Briles DE;

XX WPI; 2004-192068/18.

XX Treating Streptococcus pneumoniae infection in a subject lacking a  
 PT functional spleen comprises administering an antibody that recognizes  
 PT pneumococcal surface protein A (PspA) or its binding portion.

XX Claim 17; SEQ ID NO 3; 41pp; English.

XX The present invention relates to treating Streptococcus pneumoniae  
 CC infection in a subject lacking a functional spleen comprises  
 CC administering an antibody that recognizes pneumococcal surface protein A  
 CC (PspA) or its binding portion. The method is useful for treating or  
 CC preventing Streptococcus pneumoniae infection in a subject lacking a  
 CC functional spleen. The disease-associated injury is especially due to  
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell  
 CC anemia or Hodgkin's disease. The present sequence represents PspA  
 CC molecule from the EF3296 strain of Streptococcus pneumoniae.

XX Sequence 526 AA;

Query Match 92.7%; Score 468; DB 8; Length 526;  
 Best Local Similarity 96.2%; Pred. No. 8.6e-35;  
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKGTODELDKEA-EAELOKKADELPNKVADLEKEISNLEILG 58  
 346 LAKKQTELEKLLDLDPEKGTODELDKEAEAELOKKADELPNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEE 102  
 |||||  
 Db 406 GADSEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEE 449

RESULT 5  
 ABU00449  
 ID ABU00449 standard; protein; 744 AA.

XX ABU00449;

XX

DT 23-OCT-2003 (revised)  
 DT 11-FEB-2003 (first entry)  
 DE S. pneumoniae type 4 strain protein from coding region #16.  
 XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
 KW antinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
 KW gene therapy; vaccine.  
 XX Streptococcus pneumoniae; type 4 strain.  
 OS WO200277021-A2.  
 PN 03-OCT-2002.  
 PD 27-MAR-2002; 2002WO-IB002163.  
 PF 27-MAR-2001; 2001GB-00007658.  
 PR (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX Masignani V, Tettelin H, Fraser C;  
 PI WPI; 2003-040579/03.  
 DR N-PSDB; ABX05728.  
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
 PT useful as medicaments for treating or preventing a disease or infection  
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or  
 PT ear infection.

Claim 1; SEQ ID NO 32; 56pp; English.

XX The invention relates to a protein comprising or having at least 50%  
 CC identity to any of the 2469 amino acid sequences, identified in the  
 CC specification (available on a computer readable format), or its fragment,  
 CC expressed from 2469 identified DNA coding regions from the  
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as  
 CC ABS56454. Also included are an antibody which binds one of the proteins,  
 CC treating a patient by administering the protein, DNA or antibody (in a  
 CC composition), a kit comprising first and second primers, which are the  
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a  
 CC sequence not defined in the specification, for amplifying a target  
 CC sequence contained within a Streptococcus nucleic acid sequence, where  
 CC the first primer is substantially complementary to the target sequence  
 CC and the second primer is substantially complementary to the complement of  
 CC the target sequence, and where the parts of the primers having  
 CC substantial complementarity define the termini of the target sequence to  
 CC be amplified, assay comprising contacting a test compound with the  
 CC protein, and determining whether the test compound binds to the protein  
 CC and a Streptococcus pneumoniae bacterium, where one or more genes  
 CC encoding the proteins has been rendered inactive. The proteins, nucleic  
 CC acid molecules, antibody and compositions are useful as medicaments for  
 CC treating or preventing a disease or infection due to streptococcus  
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis  
 CC media or ear infection. They are also useful in developing vaccines,  
 CC diagnostics and antibiotics. The methods are useful for identifying  
 CC immunodominant proteins. The present sequence is one of the 2469 proteins  
 CC expressed by the identified coding regions from the genomic sequence.  
 CC Note: the sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 23-OCT-2003 to  
 CC standardise OS field)

XX Sequence 744 AA;

Query Match 92.7%; Score 468; DB 6; Length 744;  
 Best Local Similarity 96.2%; Pred. No. 1.3e-34;  
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKGTODELDKEA-EAELOKKADELPNKVADLEKEISNLEILG 58  
 |||||

Db 346 LAKKQTELEKLLDLSLDPGKGTQDELDKAEAEALDKKADLQNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEEE 102  
 |||||

Db 406 GADSEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEEE 449

RESULT 6  
 ADM92054  
 ID ADM92054 standard; protein; 744 AA.  
 XX  
 AC ADM92054;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE S pneumoniae antigenic protein sequence SeqID251.  
 XX  
 KW antibacterial; gene therapy; Streptococcus pneumoniae infection;  
 KW antigenic.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO2004020609-A2.  
 XX  
 PD 11-MAR-2004.  
 XX  
 PF 02-SEP-2003; 2003WO-US027401.  
 XX  
 PR 30-AUG-2002; 2002US-0407082P.  
 XX  
 PA (TUFT ) UNIV TUFTS.  
 XX  
 PI Camilli A, Hava DL;  
 XX  
 DR WPI; 2004-239189/22.  
 DR N-PSDB; ADM91817.  
 XX  
 PT New Streptococcus pneumoniae nucleic acid molecules, useful for  
 PT diagnosing, treating and preventing active infections of Streptococcus  
 PT pneumoniae.  
 XX  
 PS Claim 27; SEQ ID NO 251; 123pp; English.  
 XX  
 CC This invention relates to novel isolated Streptococcus pneumoniae nucleic  
 CC acid molecules and the antigenic polypeptides encoded by them. The  
 CC invention may be useful for the production of compounds with an  
 CC antibacterial activity or for gene therapy. The nucleic acid molecules,  
 CC compositions and methods disclosed are useful for treating Streptococcus  
 CC pneumoniae infection. The present sequence is that of an S pneumoniae  
 CC protein of the invention.  
 XX  
 SQ Sequence 744 AA;

Query Match 92.7%; Score 468; DB 8; Length 744;  
 Best Local Similarity 96.2%; Pred. No. 1.3e-34;  
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKGTQDELDKAEAEALDKKADLQNKVADLEKEISNLEILG 58  
 |||||

Db 346 LAKKQTELEKLLDLSLDPGKGTQDELDKAEAEALDKKADLQNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEEE 102  
 |||||

Db 406 GADSEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEEE 449

RESULT 7  
 AAY81652  
 ID AAY81652 standard; protein; 745 AA.  
 XX  
 AC AAY81652;  
 XX  
 DT 24-MAY-2000 (first entry)

XX Streptococcus pneumoniae protein sequence ID301.  
 DE  
 XX Streptococcus pneumoniae; vaccine; screening; protein antigen;  
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;  
 KW pneumococcal disease.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO200006737-A2.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 27-JUL-1999; 99WO-GB002451.  
 XX  
 PR 27-JUL-1998; 98GB-00016337.  
 PR 19-MAR-1999; 99US-0125164P.  
 XX  
 PA (MICR-) MICROBIAL TECHNIQS LTD.  
 XX  
 PI Gilbert CFG, Hanabro PM;  
 XX  
 DR WPI; 2000-195300/17.  
 XX  
 PT New Streptococcal protein, useful as a vaccine, for diagnosis of  
 PT pneumococcal diseases and for screening agents capable of antagonizing or  
 PT inhibiting expression of the protein.  
 XX  
 PS Claim 2; Page 95; 109pp; English.  
 XX  
 CC AAY81501 to AAY81679 represent specifically claimed protein sequences  
 CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent  
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.  
 CC The sequences have antibacterial and antiinflammatory properties. The  
 CC protein sequences, and fragments of them, are useful as immunogens and/or  
 CC antigens. The nucleotide sequences can be used in vaccines and in  
 CC diagnostic assays. The proteins and nucleotides can be useful for the  
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also  
 CC useful for screening an agent capable of antagonising, inhibiting or  
 CC interfering with the function or expression of the proteins in which the  
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection  
 CC and meningitis. AAA05591 to AAA05614 represent primers used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 745 AA;

Query Match 92.7%; Score 468; DB 3; Length 745;  
 Best Local Similarity 96.2%; Pred. No. 1.3e-34;  
 Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKGTQDELDKAEAEALDKKADLQNKVADLEKEISNLEILG 58  
 |||||

Db 346 LAKKQTELEKLLDLSLDPGKGTQDELDKAEAEALDKKADLQNKVADLEKEISNLEILG 405

QY 59 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEEE 102  
 |||||

Db 406 GADSEDDTAALQNKLTAKKAELEKTQKELDAALNELGPDGDEEE 449

RESULT 8  
 AAW14567  
 ID AAW14567 standard; protein; 213 AA.  
 XX  
 AC AAW14567;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX

OS Streptococcus pneumoniae; strain Bg8090.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2  
 FT /note= "unidentified amino acid"  
 XX  
 PN XX  
 XX WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 XX 16-SEP-1996; 96WO-US014819.  
 XX  
 PF 15-SEP-1995; 95US-00529055.  
 XX  
 PR (UABR-) UAB RES FOUND.  
 XX  
 PA Briles DB, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 XX Hollingshead S, Tart R, Brooks-Walter A;  
 PI WPI; 1997-202002/18.  
 DR  
 XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 PT  
 XX Example 6; Fig 13; 296pp; English.  
 XX  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8090.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 213 AA;  
 Query Match 92.5%; Score 467; DB 2; Length 213;  
 Best Local Similarity 95.3%; Pred. No. 3.7e-35;  
 Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
 QY 1 LAKKQTELEKLLD-LDPEGKTQDELQKE-AEALDKKADLPNKVADLEKEISNLSILLG 58  
 DB 59 LAKKQTELEKLLDLDPEGKTQDELQKEAAEALDKKADLPNKVADLEKEISNLSILLG 118  
 QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 102  
 DB 119 GADPEDDTAALPNKLTAKKAEFEKTFKELDAALNELGPDGDEEE 162  
 RESULT 9  
 AAW61217  
 ID AAW61217 standard; protein; 641 AA.  
 XX  
 AC AAW61217;  
 XX  
 XX 02-OCT-1998 (first entry)  
 XX  
 DE Streptococcus pneumoniae SP0092 protein.  
 XX  
 XX Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;  
 KW detection; pneumonia; otitis media; meningitis.  
 XX  
 XX Streptococcus pneumoniae.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 306  
 FT /label= unknown  
 FT

FT  
 XX /note= "encoded by NCT"  
 PN WO9818930-A2.  
 XX  
 PD 07-MAY-1998.  
 XX  
 XX 30-OCT-1997; 97WO-US019422.  
 PF  
 XX 31-OCT-1996; 96US-0029960P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Kunesh CA, Choi GH, Johnson LS, Hromockyj A;  
 PI WPI; 1998-272224/24.  
 XX  
 DR N-PSDB; AAV27403.  
 DR  
 XX Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae  
 PT - or their epitope-containing fragments, useful in protective or  
 PT therapeutic vaccines, and for diagnosis.  
 PT  
 XX Claim 11; Page 82; 118pp; English.  
 PS  
 XX The present sequence represents a protein from Streptococcus pneumoniae.  
 CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein  
 CC can be useful in vaccines for inducing protective antibodies against  
 CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.  
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
 CC are used to detect Streptococcus infection (by usual hybridisation or  
 CC amplification methods), also for isolating Streptococcus genes or their  
 CC allelic variants. The protein can be used similarly to detect specific  
 CC antibodies in standard immunoassays, especially for diagnosing or  
 CC monitoring infections. Antibodies which bind the protein are used to  
 CC detect corresponding antigens, to purify the protein and for passive  
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000  
 CC (especially 10-300) mu g/ml per dose  
 XX  
 SQ Sequence 641 AA;  
 Query Match 91.9%; Score 464; DB 2; Length 641;  
 Best Local Similarity 95.2%; Pred. No. 2.5e-34;  
 Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
 QY 1 LAKKQTELEKLLD-LDPEGKTQDELQKEA-EAELDKKADLPNKVADLEKEISNLSILLG 58  
 DB 243 LAKKQTELEKLLDLDPEGKTQDELQKEAAEALDKKADLPNKVADLEKEISNLSILLG 302  
 QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 102  
 DB 303 GADXEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 346  
 RESULT 10  
 ABP54636  
 ID ABP54636 standard; protein; 641 AA.  
 XX  
 AC ABP54636;  
 XX  
 XX 04-SEP-2002 (first entry)  
 DT  
 XX  
 DE S. pneumoniae SP092 protein sequence SEQ ID NO:160.  
 XX  
 XX Streptococcus pneumoniae; epitope; vaccine; antigenic protein;  
 KW antibacterial; Streptococcal infection; detection.  
 XX  
 XX Streptococcus pneumoniae.  
 OS  
 XX US2002061545-A1.  
 PN  
 XX 23-MAY-2002.  
 PD  
 XX 22-JAN-2001; 2001US-00765272.  
 PF

```

XX 30-OCT-1997; 97US-00961083.
XX (CHOI/) CHOI G H.
XX (KUNS/) KUNSCH C A.
XX (BARA/) BARASH S C.
XX (DILL/) DILLON P J.
XX (DOUG/) DOUGHERTY B.
XX (FANN/) FANNON M R.
XX (ROSE/) ROSEN C A.
XX
XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
XX Rosen CA;
XX
XX WPI: 2002-479261/51.
XX N-PSDB; ABQ84871.
XX
XX New Streptococcus pneumoniae antigens, useful for detecting Streptococcus
XX and for preventing or attenuating disease caused by Streptococcus
XX infection.
XX
XX Claim 11; Page 43; 70pp; English.
XX
XX ABQ84792 to ABQ84904 represents nucleic acids which encode the
XX Streptococcus pneumoniae antigens given in ABP54557 to ABP54669. The S.
XX pneumoniae antigens have antibacterial activity and can be used in
XX vaccines. The S. pneumoniae antigens can also be used to prevent or
XX attenuate a Streptococcal infection in an animal. The polynucleotides
XX encoding the S. pneumoniae antigens can be used to detect Streptococcus
XX nucleic acids. ABQ84905 to ABQ85130 represent primers used in the cloning
XX of S. pneumoniae ORFs (open reading frames) which are used in an example
XX from the present invention
XX
XX Sequence 641 AA;
XX
XX Query Match 91.9%; Score 464; DB 5; Length 641;
XX Best Local Similarity 95.2%; Pred. No. 2.5e-34;
XX Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
XX
XX 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEISNLEILLG 58
XX 243 LAKKQTELEKLLDLSLDPEGKTQDELKKEAEAEALDKKADLPKNKADLEKEISNLEILLG 302
XX
XX 59 GADSEDDTAALPNKATKKALEKTKQKELDAALNELGPDGDEEE 102
XX 303 GADXEDDTAALQNKLATKKALEKTKQKELDAALNELGPDGDEEE 346
XX
XX
XX RESULT 11
XX ADC45241
XX ID ADC45241 standard; protein; 641 AA.
XX
XX AC ADC45241;
XX
XX 18-DEC-2003 (first entry)
XX
XX S. pneumoniae antigenic protein SP092.
XX
XX Antigen; bacterial infection; vaccine; pneumonia; antibacterial.
XX
XX Streptococcus pneumoniae.
XX
XX US6573082-B1.
XX
XX 03-JUN-2003.
XX
XX 28-MAR-2000; 2000US-00536784.
XX
XX 31-OCT-1996; 96US-0029960P.
XX 30-OCT-1997; 97US-00961083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX

```

```

XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
XX Rosen CA;
XX
XX WPI: 2003-764574/72.
XX N-PSDB; ADC45240.
XX
XX Novel polynucleotide encoding Streptococcus pneumoniae polypeptides
XX useful for producing vaccines for prevention or attenuation of infection
XX by Streptococcus pneumoniae.
XX
XX Example 1; SEQ ID NO 160; 58pp; English.
XX
XX The invention relates to an isolated polynucleotide consisting of a
XX Streptococcus pneumoniae nucleic acid (appearing as ADC45122 and encoding
XX SP028) one of 113 disclosed nucleic acids encoding 113 S. pneumoniae
XX antigens. Also included are making a recombinant vector by inserting the
XX nucleic acid into a vector, an isolated polynucleotide consisting of at
XX least 50 or 100 contiguous nucleotides of the SP028 nucleic acid, and a
XX recombinant host cell comprising the SP028 polynucleotide. The nucleic
XX acids are useful as DNA vaccine against Streptococcus pneumoniae
XX infection (e.g. pneumonia). Nucleic acids derived from the S. pneumoniae
XX antigen nucleic acids are useful as probes for use in diagnostic methods
XX for detecting S. pneumoniae gene expression. The present sequence
XX represents an S. pneumoniae antigenic protein.
XX
XX Sequence 641 AA;
XX
XX Query Match 91.9%; Score 464; DB 7; Length 641;
XX Best Local Similarity 95.2%; Pred. No. 2.5e-34;
XX Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
XX
XX 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEISNLEILLG 58
XX 243 LAKKQTELEKLLDLSLDPEGKTQDELKKEAEAEALDKKADLPKNKADLEKEISNLEILLG 302
XX
XX 59 GADSEDDTAALPNKATKKALEKTKQKELDAALNELGPDGDEEE 102
XX 303 GADXEDDTAALQNKLATKKALEKTKQKELDAALNELGPDGDEEE 346
XX
XX
XX RESULT 12
XX ABW02598
XX ID ABW02598 standard; protein; 197 AA.
XX
XX AC ABW02598;
XX
XX 12-FEB-2004 (first entry)
XX
XX Ac122c pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI: 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding
XX

```

```

PT pneumococcal surface protein A.
PS Example 6; SEQ ID NO 44; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antibodies, an
CC immunological or vaccine compositions, for eliciting antigenic, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Acl22c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 197 AA;

Query Match      91.7%; Score 463; DB 7; Length 197;
Best Local Similarity 94.2%; Pred. No. 7.8e-35;
Matches 98; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKEA-EALDKKADLPNKVADLEKEISNLEILG 58
   ||:|||||
DB 22 LAQKQTELGKLLDLDPEGKTQDELDKEAAGEALDKKADGLPNKVSDELEKEISNLEILG 81
   ||:|||||

QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102
   |||||||
DB 82 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 125
   |||||||

RESULT 13
ABW02606
ID ABW02606 standard; protein; 233 AA.
XX
AC ABW02606;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bfl019c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..233
FT /note= "Xaa = Unknown amino acid"
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PP 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
DR WPI; 2003-862841/80.
XX
CC Immunological composition for obtaining expression products used for
CC detecting the presence of Streptococcus pneumoniae or its strain.
CC comprises at least two different full length isolated gene encoding
CC pneumococcal surface protein A.
XX

```

---

```

PS Example 6; SEQ ID NO 52; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antibodies, an
CC immunological or vaccine compositions, for eliciting antigenic, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Bfl019c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 233 AA;

Query Match      87.5%; Score 442; DB 7; Length 233;
Best Local Similarity 91.3%; Pred. No. 8.3e-33;
Matches 95; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELDKEA-EALDKKADLPNKVADLEKEISNLEILG 58
   ||:|||||
DB 53 LAQKQTELGKLLDLDPEGKTQDELDKEAAGEALDKKADLPNKVADLEKEISNLEILG 112
   ||:|||||

QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102
   |||||||
DB 113 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 156
   |||||||

RESULT 14
AAW14564
ID AAW14564 standard; protein; 196 AA.
XX
AC AAW14564;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain Acl22.
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PP 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 1997-202002/18.
XX
PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
PS Example 6; Fig 13; 296pp; English.
XX
CC This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain Acl22.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 19.0469 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-18

Perfect score: 505

Sequence: 1 LAKQTELEKLLDPEGKT.....TQKELDAALNELGPDGDEE 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfilese1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 505   | 100.0       | 102    | 2     | US-08-710-749-21  |
| 2          | 505   | 100.0       | 102    | 4     | US-09-147-875A-18 |
| 3          | 484   | 95.8        | 104    | 2     | US-08-710-749-19  |
| 4          | 484   | 95.8        | 104    | 4     | US-09-147-875A-20 |
| 5          | 479   | 94.9        | 104    | 4     | US-09-147-875A-21 |
| 6          | 473   | 93.7        | 104    | 2     | US-08-710-749-20  |
| 7          | 469   | 92.9        | 213    | 4     | US-08-529-055-47  |
| 8          | 469   | 92.9        | 891    | 4     | US-08-714-741-32  |
| 9          | 464   | 91.9        | 641    | 3     | US-08-561-083-160 |
| 10         | 464   | 91.9        | 641    | 4     | US-09-536-784-160 |
| 11         | 463   | 91.7        | 137    | 4     | US-08-529-055-44  |
| 12         | 442   | 87.5        | 233    | 4     | US-08-529-055-52  |
| 13         | 369.5 | 73.2        | 80     | 2     | US-08-710-749-18  |
| 14         | 369.5 | 73.2        | 80     | 4     | US-09-147-875A-19 |
| 15         | 298   | 59.0        | 108    | 2     | US-08-710-749-26  |
| 16         | 298   | 59.0        | 108    | 4     | US-09-147-875A-23 |
| 17         | 298   | 59.0        | 211    | 4     | US-08-529-055-67  |
| 18         | 293   | 58.0        | 108    | 2     | US-08-710-749-24  |
| 19         | 293   | 58.0        | 108    | 4     | US-09-147-875A-25 |
| 20         | 293   | 58.0        | 232    | 4     | US-08-529-055-70  |
| 21         | 293   | 58.0        | 458    | 4     | US-08-529-055-73  |
| 22         | 290   | 57.4        | 108    | 4     | US-09-147-875A-24 |
| 23         | 288   | 57.0        | 106    | 4     | US-09-147-875A-22 |
| 24         | 283   | 56.0        | 212    | 4     | US-08-529-055-68  |
| 25         | 280   | 55.4        | 108    | 2     | US-08-710-749-22  |
| 26         | 280   | 55.4        | 108    | 4     | US-08-710-749-23  |
| 27         | 243   | 48.1        | 108    | 2     | US-08-710-749-25  |

28 243 48.1 108 4 US-09-147-875A-26 Sequence 26, Appl  
29 236 46.7 185 4 US-08-529-055-69 Sequence 69, Appl  
30 177.5 35.1 119 2 US-08-710-749-27 Sequence 27, Appl  
31 177.5 35.1 119 4 US-09-147-875A-27 Sequence 27, Appl  
32 177.5 35.1 215 4 US-08-529-055-43 Sequence 43, Appl  
33 126 25.0 550 4 US-09-583-110-4871 Sequence 4871, Ap  
34 126 25.0 550 4 US-09-107-433-3858 Sequence 3858, Ap  
35 126 25.0 623 4 US-08-714-741-47 Sequence 47, Appl  
36 121.5 24.1 101 2 US-08-710-749-1 Sequence 1, Appl  
37 121.5 24.1 204 4 US-08-529-055-51 Sequence 51, Appl  
38 121 24.0 605 4 US-08-714-741-46 Sequence 46, Appl  
39 120.5 23.9 101 2 US-08-710-749-2 Sequence 2, Appl  
40 119.5 23.7 289 1 US-08-072-070-4 Sequence 4, Appl  
41 119.5 23.7 289 1 US-08-469-434-4 Sequence 4, Appl  
42 119.5 23.7 289 1 US-08-214-222-4 Sequence 4, Appl  
43 119.5 23.7 289 2 US-08-467-852A-5 Sequence 5, Appl  
44 119.5 23.7 289 2 US-08-468-718-4 Sequence 4, Appl  
45 119.5 23.7 289 2 US-08-247-491A-5 Sequence 5, Appl

#### ALIGNMENTS

RESULT 1  
US-08-710-749-21  
; Sequence 21, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Brilles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-21

Query Match 100.0%; Score 505; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 2.7e-42;  
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAKQTELEKLLDPEGKTQDELQKAEAELEKKADELQKVADELEKISNLEILGGA 60  
DB 1 LAKQTELEKLLDPEGKTQDELQKAEAELEKKADELQKVADELEKISNLEILGGA 60

```
Qy 61 DSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||||
Db 61 DSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 102

RESULT 2
US-09-147-875A-18
; Sequence 18, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-18

Query Match 100.0%; Score 505; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.7e-42;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDLDPEGKTQDELDKAEAEALDKKADELPNKVADLEKEISNLEILGGA 60
      |||||||
Db 1 LAKQTELEKLLDLDPEGKTQDELDKAEAEALDKKADELPNKVADLEKEISNLEILGGA 60

Qy 61 DSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||||
Db 61 DSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 102

RESULT 3
US-08-710-749-19
; Sequence 19, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Brilles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-19

Query Match 95.8%; Score 484; DB 2; Length 104;
Best Local Similarity 98.1%; Pred. No. 3.1e-40;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 LAKQTELEKLLDLDPEGKTQDELDKAEAEALDKKADELPNKVADLEKEISNLEILG 58
      |||||||
Db 1 LAKQTELEKLLDLDPEGKTQDELDKAEAEALDKKADELPNKVADLEKEISNLEILG 60

Qy 59 GADSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||||
Db 61 GADSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 104
```

```
RESULT 4
US-09-147-875A-20
; Sequence 20, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-20

Query Match 95.8%; Score 484; DB 4; Length 104;
Best Local Similarity 98.1%; Pred. No. 3.1e-40;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 LAKQTELEKLLDLDPEGKTQDELDKAEAEALDKKADELPNKVADLEKEISNLEILG 58
      |||||||
Db 1 LAKQTELEKLLDLDPEGKTQDELDKAEAEALDKKADELPNKVADLEKEISNLEILG 60

Qy 59 GADSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 102
      |||||||
Db 61 GADSEDDTAALPNKATKAELEKTKQKELDAALNELGPDGDEEE 104
```

```
RESULT 5
US-09-147-875A-21
; Sequence 21, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-21

Query Match 94.9%; Score 479; DB 4; Length 104;
Best Local Similarity 97.1%; Pred. No. 9.6e-40;
Matches 101; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
```



|  |    |    |  |     |
|--|----|----|--|-----|
|  | Qy | 1  | LAKQTELEKLLD-IDPEGKTQDELDOKE-AEALDKKADELPNKVADLEKEISNIEILLIG   | 58  |
|  |    |    |  |     |
|  | Db | 1  | LAKQTELEKLLNLNLOPEGKTQDELDOKEAAEAELDKKADELPNKVADLEKEISNIEILLIG | 60  |
|  |    |    |  |     |
|  | Qy | 59 | GADSEDDTAALPNKLATKKAELAETKOKELDAALNEILGPDGDEEE                 | 102 |
|  |    |    |  |     |
|  | Db | 61 | GAPEDDTAALPNKLATKKAELAETKOKELDAALNEILGPDGDEEE                  | 104 |
|  |    |    |  |     |

## RESULT 6

US-08-710-749-20  
 ; Sequence 20, Application US/08710749  
 ; Patent No. 5955089  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E.  
 ; APPLICANT: Hollingshead, Susan  
 ; APPLICANT: Becker, Robert  
 ; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
 ; TITLE OF INVENTION: PROTEINS  
 ; NUMBER OF SEQUENCES: 28  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Curtis, Morris & Safford  
 ; STREET: 530 Fifth Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036

REF. 10030  
COMPUTER READABLE FORM:

C. MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/710,749  
 FILING DATE: 20-SEP-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454312-2074  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 104 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: n/a  
 TOPOLOGY: linear  
 MOLECULE TYPE: amino acid  
 US-08-710-749-20

## Query Match

|    | Query Match           | 93.7%  | Score 473;         | DB 2;     | Length 104; |
|----|-----------------------|--|--------------------|-----------|-------------|
|    | Best Local Similarity | 96.2%;   | Prod. No. 3.7e-39; |           |             |
|    | Matches 100;          | Conservative 0;  | Mismatches 2;      | Indels 2; | Gaps 2;     |
| Qy | 1                     | LAKQTELEKLLD-LDPEGKTQDDELKE-ABAEALDKKADELPNKVADLEKEISNLEILLG | 58                 |           |             |
| Db | 1                     | LAKQTELEKLLDNLDPGKTQDDELKEAAEALDKKADELPNKVADLEKEISNLEILLG    | 60                 |           |             |
| Qy | 59                    | GADSEDDTAALPNKLTATKKALEKTKQELDAAALNELPGDGDDEE                | 102                |           |             |
| Db | 61                    | GADPEDDTAALPNKLTATKKALEKTKPELDAALNELPGDGDDEE                 | 104                |           |             |

## RESULT 7

US-08-529-055-47  
; Sequence 47. Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiatlo, Edwin

APPLICANT: Yother, Janet  
 APPLICANT: Brooks-Walter, Alexis  
 TITLE OF INVENTION: Pneumococcal Genes, Portions  
 TITLE OF INVENTION: Thereof, Expression Products  
 TITLE OF INVENTION: Thereof, and Uses of Such Genes,  
 TITLE OF INVENTION: Portions and Products  
 NUMBER OF SEQUENCES: 73  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Curtis, Morris & Safford, P.C.  
 STREET: 530 Fifth Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/529,055  
 FILING DATE: 15-SEP-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454312-2400  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 INFORMATION FOR SEQ ID NO: 47:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 213 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-529-055-47

Query Match 92.9%; Score 469; DB 4; Length 213;  
Best Local Similarity 95.2%; Pred. No. 2.2e-38;  
Matches 99; Conservative 0; Mismatches 3; Indels

|    |     |   |     |
|----|-----|---|-----|
| Qy | 1   | LAKQTELEKLLD-LDPGKQTQDELDK3-ABAE1DKKADELPNKVAOLEKEISNLE1LLG | 58  |
| Db | 59  | LAKQTELEKLLDNLDPGKQTQDELDK3ABAE1DKKADELPNKVAOLEKEISNLE1LLG  | 118 |
| Qy | 59  | GADSEDDTAALPNK1LAKKAELEKTKQELDKALNALGPDGDEEE                | 102 |
| Db | 119 | GADPEDDTAALPNK1LAKKAELEKTKQELDKALNALGPDGDEEE                | 162 |

## RESULT 8

RESULT 8  
US-08-714-741-32  
Sequence 32, Application US/08714741  
Patent No. 6500613  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J.  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
PORTIONS THEREOF, AND USES OF SUCH GENES,  
EXPRESSION PRODUCTS THEREFROM  
TITLE OF INVENTION: PORTIONS AND PRODUCTS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York

```
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32

Query Match          92.9%; Score 469; DB 4; Length 8991;
Best Local Similarity 95.2%; Pred. No. 2.3e-36;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 58
DB 4625 LAKKQTELEKLLDLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 4684

QY 59 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 102
DB 4685 GADPEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 4728

RESULT 9
US-08-961-083-160
; Sequence 160, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-536-784-160

Query Match          91.9%; Score 464; DB 4; Length 641;
Best Local Similarity 95.2%; Pred. No. 2.7e-37;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 58
DB 243 LAKKQTELEKLLDLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 302

QY 59 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 102
DB 303 GADXEDDTAALQNKLTAKAELEKTQKELDAALNELGPDGDEEE 346

RESULT 10
US-09-536-784-160
; Sequence 160, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-536-784-160

Query Match          91.9%; Score 464; DB 4; Length 641;
Best Local Similarity 95.2%; Pred. No. 2.7e-37;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 58
```

```
|||||
Db 243 LAKKQTELEKLSLDPEGKTQDELQKAEAEALDKKADLPNKVADLEKEISNLEILG 302
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 102
Db 303 GADXEDDTAALQNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 346

RESULT 11
US-08-529-055-44
; Sequence 44, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yotter, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-44

Query Match 91.7%; Score 463; DB 4; Length 197;
Best Local Similarity 94.2%; Pred. No. 7.7e-38;
Matches 98; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLD-LDPEGKTQDELQKAEAEALDKKADLPNKVADLEKEISNLEILG 59
Db 22 LAKQTELEKLSLDPEGKTQDELQKAEAEALDKKADLPNKVADLEKEISNLEILG 81

QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 102
Db 82 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 125

RESULT 12
US-08-529-055-52
; Sequence 52, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yotter, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-44

Query Match 91.7%; Score 463; DB 4; Length 197;
Best Local Similarity 94.2%; Pred. No. 7.7e-38;
Matches 98; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLD-LDPEGKTQDELQKAEAEALDKKADLPNKVADLEKEISNLEILG 59
Db 22 LAKQTELEKLSLDPEGKTQDELQKAEAEALDKKADLPNKVADLEKEISNLEILG 81

QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 102
Db 82 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 125

RESULT 13
US-08-710-749-18
; Sequence 18, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
```

```
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yotter, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-52

Query Match 87.5%; Score 442; DB 4; Length 233;
Best Local Similarity 91.3%; Pred. No. 1.1e-35;
Matches 95; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLD-LDPEGKTQDELQKAEAEALDKKADLPNKVADLEKEISNLEILG 58
Db 53 LAKQTELEKLSLDPEGKTQDELQKAEAEALDKKADLPNKVADLEKEISNLEILG 112

QY 59 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 102
Db 113 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNEILGPDGDEE 156

RESULT 13
US-08-710-749-18
; Sequence 18, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 65.1551 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-18  
Perfect score: 505  
Sequence: 1 LAKKQTELEKLLDPEGKT.....TKELDAAALNELGPDGDEEE 102

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No.        | Score | Query Match | Length | ID | Description       |
|-------------------|-------|-------------|--------|----|-------------------|
| 1                 | 505   | 100.0       | 102    | 15 | US-10-674-755-18  |
| 2                 | 484   | 95.8        | 104    | 15 | US-10-674-755-20  |
| 3                 | 479   | 94.9        | 104    | 15 | US-10-674-755-21  |
| 4                 | 469   | 92.9        | 213    | 15 | US-10-299-636-62  |
| 5                 | 468   | 92.7        | 744    | 10 | US-09-769-787-184 |
| 6                 | 468   | 92.7        | 744    | 17 | US-10-472-928-32  |
| 7                 | 464   | 91.9        | 641    | 9  | US-09-765-272-160 |
| 8                 | 463   | 91.7        | 197    | 15 | US-10-299-636-59  |
| 9                 | 442   | 87.5        | 233    | 15 | US-10-299-636-67  |
| 10                | 369.5 | 73.2        | 80     | 15 | US-10-674-755-19  |
| 11                | 298   | 59.0        | 108    | 15 | US-10-674-755-23  |
| Sequence 18, Appl |       |             |        |    |                   |
| Sequence 20, Appl |       |             |        |    |                   |
| Sequence 21, Appl |       |             |        |    |                   |
| Sequence 62, Appl |       |             |        |    |                   |
| Sequence 184, App |       |             |        |    |                   |
| Sequence 32, Appl |       |             |        |    |                   |
| Sequence 160, App |       |             |        |    |                   |
| Sequence 59, Appl |       |             |        |    |                   |
| Sequence 67, Appl |       |             |        |    |                   |
| Sequence 19, Appl |       |             |        |    |                   |
| Sequence 23, Appl |       |             |        |    |                   |

|    |       |      |     |    |                      |                   |
|----|-------|------|-----|----|----------------------|-------------------|
| 12 | 298   | 59.0 | 211 | 15 | US-10-299-636-82     | Sequence 82, Appl |
| 13 | 293   | 58.0 | 108 | 15 | US-10-674-755-25     | Sequence 25, Appl |
| 14 | 293   | 58.0 | 232 | 15 | US-10-299-636-85     | Sequence 85, Appl |
| 15 | 293   | 58.0 | 275 | 16 | US-10-414-532-1      | Sequence 1, Appl  |
| 16 | 293   | 58.0 | 458 | 15 | US-10-299-636-88     | Sequence 88, Appl |
| 17 | 293   | 58.0 | 653 | 16 | US-10-414-532-26     | Sequence 26, Appl |
| 18 | 290   | 57.4 | 108 | 15 | US-10-674-755-24     | Sequence 24, Appl |
| 19 | 288   | 57.0 | 106 | 15 | US-10-674-755-22     | Sequence 22, Appl |
| 20 | 283   | 56.0 | 212 | 15 | US-10-299-636-83     | Sequence 83, Appl |
| 21 | 253   | 50.1 | 459 | 16 | US-10-702-305A-18    | Sequence 18, Appl |
| 22 | 243   | 48.1 | 108 | 15 | US-10-674-755-26     | Sequence 84, Appl |
| 23 | 236   | 46.7 | 185 | 15 | US-10-299-636-84     | Sequence 86, Appl |
| 24 | 178   | 35.2 | 487 | 16 | US-10-414-532-34     | Sequence 34, Appl |
| 25 | 178   | 35.2 | 487 | 16 | US-10-414-533-21     | Sequence 21, Appl |
| 26 | 178   | 35.2 | 524 | 16 | US-10-414-532-28     | Sequence 28, Appl |
| 27 | 177.5 | 35.1 | 119 | 15 | US-10-674-755-27     | Sequence 27, Appl |
| 28 | 177.5 | 35.1 | 215 | 15 | US-10-299-636-58     | Sequence 58, Appl |
| 29 | 165.5 | 32.8 | 290 | 16 | US-10-414-532-65     | Sequence 65, Appl |
| 30 | 160   | 31.7 | 230 | 16 | US-10-414-532-32     | Sequence 32, Appl |
| 31 | 160   | 31.7 | 230 | 16 | US-10-414-533-19     | Sequence 19, Appl |
| 32 | 121.5 | 24.1 | 204 | 15 | US-10-299-636-66     | Sequence 66, Appl |
| 33 | 119   | 23.6 | 354 | 15 | US-10-299-636-105    | Sequence 105, App |
| 34 | 119   | 23.6 | 588 | 15 | US-10-299-636-96     | Sequence 96, Appl |
| 35 | 119   | 23.6 | 619 | 10 | US-09-882-774-1      | Sequence 1, Appl  |
| 36 | 119   | 23.6 | 619 | 15 | US-10-282-122A-73702 | Sequence 73702, A |
| 37 | 119   | 23.6 | 619 | 16 | US-10-414-532-72     | Sequence 72, Appl |
| 38 | 118   | 23.4 | 141 | 14 | US-10-254-995-2      | Sequence 2, Appl  |
| 39 | 118   | 23.4 | 589 | 9  | US-09-748-875-14     | Sequence 14, Appl |
| 40 | 118   | 23.4 | 589 | 10 | US-09-298-523B-14    | Sequence 14, Appl |
| 41 | 118   | 23.4 | 589 | 15 | US-10-299-636-97     | Sequence 97, Appl |
| 42 | 118   | 23.4 | 643 | 15 | US-10-299-636-95     | Sequence 95, Appl |
| 43 | 118   | 23.4 | 670 | 9  | US-09-748-875-63     | Sequence 63, Appl |
| 44 | 118   | 23.4 | 670 | 10 | US-09-298-523B-63    | Sequence 63, Appl |
| 45 | 118   | 23.4 | 690 | 9  | US-09-748-875-61     | Sequence 61, Appl |

ALIGNMENTS

RESULT 1  
US-10-674-755-18  
; Sequence 18, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 102  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-18

Query Match 100.0%; Score 505; DB 15; Length 102;  
Best Local Similarity 100.0%; Pred. No. 7.8e-36;  
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |    |   |     |
|----|----|---|-----|
| QY | 1  | LAKKQTELEKLLDPEGKTQDELKAEAEALDKKADELPNKVADLEKISNLEILGGA | 60  |
| Db | 1  | LAKKQTELEKLLDPEGKTQDELKAEAEALDKKADELPNKVADLEKISNLEILGGA | 60  |
| QY | 61 | DSEDDTAALPNKATKKALEKTKQELDAALNELGPDGDEEE                | 102 |
| Db | 61 | DSEDDTAALPNKATKKALEKTKQELDAALNELGPDGDEEE                | 102 |

RESULT 2



TELECOMMUNICATION INFORMATION:





1 LAKKOTELEKLL-DLPPEGKTODELDK-EEAELDKADELPNKVADLEKEISNLEILG 58

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 10.2 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-18  
Perfect score: 505  
Sequence: 1 LAKKQTELEKLLDLPDGKT.....TKQELDAALNELGPDGDEEE 102

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 468   | 92.7        | 744    | 2 P95013 | pneumococcal surfa |
| 2          | 119   | 23.6        | 619    | 2 A97887 | surface protein ps |
| 3          | 119   | 23.6        | 619    | 2 A41971 | surface protein ps |
| 4          | 106.5 | 21.1        | 924    | 2 S06117 | myosin heavy chain |
| 5          | 106.5 | 21.1        | 2007   | 1 B43402 | myosin heavy chain |
| 6          | 105   | 20.8        | 629    | 2 T44607 | hypothetical prote |
| 7          | 102   | 20.2        | 281    | 2 F75216 | hypothetical prote |
| 8          | 100   | 19.8        | 233    | 2 S70531 | bbk2.11 protein pr |
| 9          | 99.5  | 19.7        | 284    | 2 S23470 | beta-tropomyosin - |
| 10         | 99.5  | 19.7        | 1976   | 2 A59252 | myosin heavy chain |
| 11         | 99    | 19.6        | 415    | 2 S35760 | fcra protein precu |
| 12         | 98.5  | 19.5        | 388    | 2 A46173 | Mrp4 protein - Str |
| 13         | 98    | 19.4        | 284    | 2 A45488 | body-wall muscle t |
| 14         | 97.5  | 19.3        | 770    | 2 S56905 | probable RNA helic |
| 15         | 97    | 19.2        | 281    | 2 A34787 | tropomyosin 1 alph |
| 16         | 97    | 19.2        | 284    | 1 TMRA8  | tropomyosin alpha  |
| 17         | 97    | 19.2        | 284    | 2 A39816 | tropomyosin 2, fib |
| 18         | 97    | 19.2        | 284    | 2 B27407 | tropomyosin alpha  |
| 19         | 97    | 19.2        | 284    | 2 A27674 | tropomyosin 3, fib |
| 20         | 97    | 19.2        | 284    | 2 A25825 | tropomyosin alpha  |
| 21         | 97    | 19.2        | 284    | 2 A60597 | tropomyosin 2, fib |
| 22         | 97    | 19.2        | 284    | 2 B39816 | tropomyosin 3, fib |
| 23         | 97    | 19.2        | 516    | 2 B84709 | hypothetical prote |
| 24         | 97    | 19.2        | 864    | 2 B90395 | purine NTPase (imp |
| 25         | 97    | 19.2        | 2139   | 2 T18296 | myosin heavy chain |
| 26         | 96.5  | 19.1        | 397    | 2 H86754 | prophage pi2 prote |
| 27         | 96.5  | 19.1        | 1932   | 2 A47297 | myosin heavy chain |
| 28         | 96    | 19.0        | 284    | 2 S24972 | tropomyosin alpha, |
| 29         | 96    | 19.0        | 284    | 2 S19691 | tropomyosin alpha, |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 96   | 19.0 | 1063 | 2 T18255 | cytoskeleton assem |
| 31 | 96   | 19.0 | 1837 | 2 T41023 | probable nuclear p |
| 32 | 96   | 19.0 | 1972 | 1 A41604 | myosin heavy chain |
| 33 | 95.5 | 18.9 | 388  | 2 S52536 | fcra 15 protein -  |
| 34 | 95.5 | 18.9 | 405  | 2 A33939 | Fc gamma (IgG) rec |
| 35 | 95.5 | 18.9 | 1509 | 1 A27224 | myosin heavy chain |
| 36 | 95   | 18.8 | 280  | 2 A22165 | tropomyosin alpha  |
| 37 | 95   | 18.8 | 1169 | 2 A64505 | PI15 homolog - Met |
| 38 | 95   | 18.8 | 1177 | 2 B75150 | chromosome segrega |
| 39 | 94.5 | 18.7 | 473  | 2 F70031 | cell wall-binding  |
| 40 | 94.5 | 18.7 | 880  | 2 F75103 | conserved hypoteth |
| 41 | 94   | 18.6 | 308  | 2 T08796 | tropomyosin - huma |
| 42 | 94   | 18.6 | 385  | 2 T20410 | hypothetical prote |
| 43 | 93.5 | 18.5 | 1175 | 2 D35815 | myosin heavy chain |
| 44 | 93.5 | 18.5 | 1201 | 2 B35815 | myosin heavy chain |
| 45 | 93.5 | 18.5 | 1475 | 2 T33318 | hypothetical prote |

ALIGNMENTS

RESULT 1

P95013  
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: F95013  
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I.  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: F95013  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-744 <KUR>  
A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:B  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0117

Query Match 92.7%; Score 468; DB 2; Length 744;  
Best Local Similarity 96.2%; Pred. No. 2.1e-25;  
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

|    |     |  |
|----|-----|--|
| QY | 1   | LAKKQTELEKLLD-LDPEGKTQDELDKA-EAELDKKADELPKNKVADLEKEISNLEILLG 58    |
| DB | 346 | LAKKQTELEKLLDSDLPDEGKTQDELDKAEAEAEELDKKADELQNKVADLEKEISNLEILLG 405 |
| QY | 59  | GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEEE 102                   |
| DB | 406 | GADSEDDTAALQNKLTAKKAELEKTKQELDAALNELGPDGDEEE 449                   |

RESULT 2

A97887  
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: A97887  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-619 <KUR>

[illegible]



Db 89 NRRIQVVEELDRA--QERLATALQKLEETEKAVID 121

RESULT 10

A59252

myosin heavy chain, nonmuscle, form IIB - human

N;Alternate names: myosin type 10; NMHC-B

N;Contains: MYOSIN ATPase (BC 3.6.4.1)

C;Species: Homo sapiens (man)

C;Date: 19-May-2000 #sequence revision 19-May-2000 #text\_change 09-Jul-2004

C;Accession: A59252; B61231; G02055

R;Phillips, C.L.; Yamakawa, K.; Adelstein, R.S.

J. Muscle Res. Cell. Motil. 16, 379-389, 1995

A;Title: Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and analysis

A;Reference number: A59252; MUID:96025307; PMID:7499478

A;Accession: A59252

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1976 <SIM>

A;Cross-references: UNIPROT:P35580; GB:M69181; NID:G641957; PIDN:AAA99177.1; PID:G641958

A;Experimental source: clone lib Lambda Zap II adult human T-cell library; cell line Jur

A;Note: between nucleotides 1942-1943 in mRNA encoding human brain MHC-B there is an alt

R;Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Gdula, D.; Adelstein, C.

Circ. Res. 69, 530-539, 1991

A;Title: Human nonmuscle myosin heavy chains are encoded by two genes located on differe

A;Reference number: A61231; MUID:913116803; PMID:1860190

A;Accession: B61231

A;Molecule type: mRNA

A;Residues: 63-237, 'K', 239-664, 'L', 666-722 <SI2>

A;Cross-references: GB:M69181; NID:G641957

R;Weir, L.

submitted to the EMBL Data Library, August 1995

A;Reference number: H00753

A;Accession: G02055

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-81 <WEI>

A;Cross-references: EMBL:U34304; NID:g1143217; PIDN:AAA84880.1; PID:g1143218

C;Genetics:

A;Gene: GDB:MVH10

A;Cross-references: GDB:127350; GDB:G00-127-350; OMIM:160776

A;Map position: 17p13-17p13

C;Superfamily: myosin heavy chain; myosin motor domain homology

C;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotid

F;88-771/Domain: myosin motor domain homology <NMO>

F;178-185/Region: nucleotide-binding motif A (P-loop)

F;559-572/Region: actin binding #status predicted

F;633-647/Region: actin binding #status predicted

F;129/Modified site: N6,N6-trimethyllysine (Lys) #status predicted

F;184/Binding site: ATP (Lys) #status predicted

F;701,711/Active site: Cys #status predicted

Query Match 19.7%; Score 99.5; DB 2; Length 1976;

Best Local Similarity 28.4%; Pred. No. 25;

Matches 33; Conservative 24; Mismatches 36; Indels 23; Gaps 4;

Qy 1 LAKQTELEKLLD-----LDPEGTQDELKAEAELEKKADELPNKVADLEKISNLEI 55

Db 1025 LAKIRNQEVWISDLERLKKKEETRELEK-AKKLDGTTQDQIAELQAIQDELK 1083

Qy 56 LLGGADSE-----DDTAALPNKLTATKAELEKTKELDAALNELGPDGDEE 102

Db 1084 QLAKEEELQALARGDDETLHKNAL-----KVVRELAQIAELQEDFESEK 1131

RESULT 11

S35760

fcrA protein precursor - Streptococcus pyogenes

C;Species: Streptococcus pyogenes

C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text\_change 09-Jul-2004

C;Accession: S35760; A42711

R;Podbielski, A.

submitted to the EMBL Data Library, November 1992

A;Reference number: S35760

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-415 <FOD>

A;Cross-references: UNIPROT:Q54859; EMBL:X69324; NID:g311759; PIDN:CAA49165.1; PID:g3117

R;Haanes, E.J.; Heath, D.G.; Cleary, P.P.

J. Bacteriol. 174, 4967-4976, 1992

A;Title: Architecture of the vir regulons of group A streptococci parallels opacity fact

A;Reference number: A42711; MUID:92332431; PMID:1385809

A;Accession: A42711

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 343-415 <HAA>

A;Cross-references: GB:M86806; NID:g153630; PIDN:AAA26887.1; PID:g153631

A;Experimental source: strain CS101, OF+

A;Note: sequence extracted from NCBI backbone (NCBIN:108942, NCBI:P:108945)

C;Superfamily: M5 protein

Query Match 19.6%; Score 99; DB 2; Length 415;

Best Local Similarity 30.2%; Pred. No. 5.4;

Matches 39; Conservative 15; Mismatches 47; Indels 28; Gaps 4;

Qy 1 LAKQTELEKLLDLDPEGTQDELKAEAE-----ELDKKAEDELPNKVAD 45

Db 188 LEAKNAEIEDLKQD-ASKTEETIANLQSEAAITLENLLGSAKHELTDLQAKLDTATAEKAK 246

Qy 46 LEKEISNLEILGGADSE-----DDTAALPNKLTATKAELEK-----TOKELDAALNE 93

Db 247 LESQETLENLLGSAKRELTDLQAKLDDANAEEKLQSQAALEKQLEATKKELADLOAK 306

Qy 94 LGPDGDEE 102

Db 307 LAATNQEK 315

RESULT 12

A46173

Mrp4 protein - Streptococcus sp. (group A)

C;Species: Streptococcus sp.

C;Date: 21-Sep-1993 #sequence revision 25-Apr-1997 #text\_change 30-May-1997

C;Accession: A46173

R;O'Toole, P.; Stenberg, L.; Rissler, M.; Lindahl, G.

Proc. Natl. Acad. Sci. U.S.A. 89, 8661-8665, 1992

A;Title: Two major classes in the M protein family in group A streptococci.

A;Reference number: A46173; MUID:92409576; PMID:1528877

A;Contents: group A

A;Accession: A46173

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-388 <OLT>

A;Note: sequence extracted from NCBI backbone (NCBIN:114063, NCBI:P:114064)

C;Superfamily: M5 protein

Query Match 19.5%; Score 98.5; DB 2; Length 388;

Best Local Similarity 31.6%; Pred. No. 5.4;

Matches 37; Conservative 15; Mismatches 40; Indels 25; Gaps 4;

Qy 1 LAKQTE-----LEKLLDLDPEGTQDELKAEAELEKKADELPNKVADLEKISNLEIL 57

Db 182 IAKLQSEAAITLENLL-----GSAKREL-----TELOAKLDTATAEKAKLESQVTTLENLL 231

Qy 58 GGADSE-----DDTAALPNKLTATKAELEKTKELDAALNELGPDGDEE 102

Db 232 GSARELTDLQAKLDAANAEEKLQSQATLEKQLEATKKELADLOAKLAATNQEK 288

RESULT 13

A45488

body-wall muscle tropomyosin - sea squirt (Ciona intestinalis)

C;Species: Ciona intestinalis

C;Date: 21-Sep-1993 #sequence revision 18-Nov-1994 #text\_change 09-Jul-2004

C;Accession: A45488



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 63.1776 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-18  
Perfect score: 505  
Sequence: 1 LAKKQTELEKLDLPDGKT.....TQKELDAALNELGPDGDEE 102

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID     | Description        |
|------------|-------|-------------|--------|--------|--------------------|
| 1          | 468   | 92.7        | 228    | Q9L5B8 | Q9L5B8 streptococc |
| 2          | 468   | 92.7        | 235    | Q9L582 | Q9L582 streptococc |
| 3          | 468   | 92.7        | 249    | Q9L5D4 | Q9L5D4 streptococc |
| 4          | 468   | 92.7        | 252    | Q9L583 | Q9L583 streptococc |
| 5          | 468   | 92.7        | 360    | Q8KQK3 | Q8KQK3 streptococc |
| 6          | 468   | 92.7        | 429    | Q9LAX7 | Q9LAX7 streptococc |
| 7          | 468   | 92.7        | 526    | Q9LAX9 | Q9LAX9 streptococc |
| 8          | 468   | 92.7        | 608    | Q8VQ55 | Q8VQ55 streptococc |
| 9          | 468   | 92.7        | 744    | Q97T39 | Q97T39 streptococc |
| 10         | 463   | 91.7        | 231    | Q9L579 | Q9L579 streptococc |
| 11         | 463   | 91.7        | 241    | Q9L580 | Q9L580 streptococc |
| 12         | 463   | 91.7        | 249    | Q9L5B7 | Q9L5B7 streptococc |
| 13         | 459   | 90.9        | 502    | Q9LAX8 | Q9LAX8 streptococc |
| 14         | 457   | 90.5        | 249    | Q9L585 | Q9L585 streptococc |
| 15         | 457   | 90.5        | 256    | Q9L590 | Q9L590 streptococc |
| 16         | 454   | 89.9        | 242    | Q9L562 | Q9L562 streptococc |
| 17         | 450   | 89.1        | 209    | Q9L593 | Q9L593 streptococc |
| 18         | 432   | 85.5        | 222    | Q9L584 | Q9L584 streptococc |
| 19         | 303   | 60.0        | 246    | Q9L5B4 | Q9L5B4 streptococc |
| 20         | 299   | 59.2        | 479    | Q9LAX2 | Q9LAX2 streptococc |
| 21         | 299   | 59.2        | 481    | Q9LAX5 | Q9LAX5 streptococc |
| 22         | 298   | 59.0        | 107    | Q8KQK2 | Q8KQK2 streptococc |
| 23         | 293   | 58.0        | 653    | Q34Q97 | Q34Q97 streptococc |
| 24         | 291   | 57.6        | 213    | Q8GNS7 | Q8GNS7 streptococc |
| 25         | 274   | 54.3        | 480    | Q9LAX3 | Q9LAX3 streptococc |
| 26         | 236.5 | 46.8        | 211    | Q8GNT0 | Q8GNT0 streptococc |
| 27         | 236.5 | 46.8        | 257    | Q9L594 | Q9L594 streptococc |
| 28         | 230.5 | 45.6        | 227    | Q9KGS0 | Q9KGS0 streptococc |
| 29         | 230.5 | 45.6        | 256    | Q9L595 | Q9L595 streptococc |
| 30         | 230.5 | 45.6        | 461    | Q9LAX6 | Q9LAX6 streptococc |
| 31         | 134   | 26.5        | 417    | Q9LAY3 | Q9LAY3 streptococc |

|    |       |      |     |   |        |                    |
|----|-------|------|-----|---|--------|--------------------|
| 32 | 128   | 25.3 | 237 | 2 | Q9L592 | Q9L592 streptococc |
| 33 | 128   | 25.3 | 395 | 2 | Q9LAY9 | Q9LAY9 streptococc |
| 34 | 126   | 25.0 | 222 | 2 | Q9L577 | Q9L577 streptococc |
| 35 | 126   | 25.0 | 262 | 2 | Q9L576 | Q9L576 streptococc |
| 36 | 126   | 25.0 | 415 | 2 | Q9LAY7 | Q9LAY7 streptococc |
| 37 | 125   | 24.8 | 393 | 2 | Q9LAZ3 | Q9LAZ3 streptococc |
| 38 | 125   | 24.8 | 416 | 2 | Q9LAY8 | Q9LAY8 streptococc |
| 39 | 124   | 24.6 | 255 | 2 | Q9L581 | Q9L581 streptococc |
| 40 | 124   | 24.6 | 255 | 2 | Q9L5B6 | Q9L5B6 streptococc |
| 41 | 123   | 24.4 | 207 | 2 | Q8GNS9 | Q8GNS9 streptococc |
| 42 | 123   | 24.4 | 225 | 2 | Q9L591 | Q9L591 streptococc |
| 43 | 122   | 24.2 | 249 | 2 | Q9L575 | Q9L575 streptococc |
| 44 | 121.5 | 24.1 | 99  | 2 | Q8KQK4 | Q8KQK4 streptococc |
| 45 | 121   | 24.0 | 246 | 2 | Q9L578 | Q9L578 streptococc |

## ALIGNMENTS

RESULT 1  
Q9L5B8 PRELIMINARY; PRT; 228 AA.  
ID Q9L5B8  
AC Q9L5B8;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=60;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones."  
RL J. Clin. Microbiol. 38:3663-3669(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=60;  
RA Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF253404; AAF67352.1;  
DR InterPro; IPR009053; Prefoldin.  
FT NON\_TER 1  
FT NON\_TER 228  
SQ SEQUENCE 228 AA; 24430 MW; E6EAA953EC54EA0F CRC64;

Query Match 92.7%; Score 468; DB 2; Length 228;  
Best Local Similarity 96.2%; Pred. No. 4.9e-24;  
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
QY 1 LAKKQTELEKLD-LDPEGKTQDELDEKA-EAELEKKADELKPKVADLEKEISNLEILLG 58  
Db 46 LAKKQTELEKLDLDPEGKTQDELDEKAEEAELEKKADELKPKVADLEKEISNLEILLG 105  
QY 59 GADSEDDTAALPNKATKAELEKTKQELDAALNELGPDGDEE 102  
Db 106 GADSEDDTAALQNKATKAELEKTKQELDAALNELGPDGDEE 149

RESULT 2  
Q9L582 PRELIMINARY; PRT; 235 AA.  
ID Q9L582  
AC Q9L582;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE PspA (Fragment).

```

GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255545; AAF68098.1; -.
FT NON_TER 1
FT NON_TER 235
SQ SEQUENCE 235 AA; 25424 MW; BFFBB48C52CA8380 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 235;
Best Local Similarity 96.2%; Pred. No. 5e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKA-EAELDKKADLPKNKVADLEKEISNLEILLG 58
Db 56 LAKKQTELEKLLDLDPEGTQDELKKA-EAELDKKADLPKNKVADLEKEISNLEILLG 115
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 116 GADSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 159
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 116 GADSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 159

RESULT 3
Q9L5D4 ID Q9L5D4 PRELIMINARY; PRT; 249 AA.
AC Q9L5D4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI95;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPI95;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252286; AAF69499.1; -.
FT HSSP; P04268; 11C2.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 26821 MW; F8EA39225CF8D43F CRC64;

Query Match 92.7%; Score 468; DB 2; Length 249;
Best Local Similarity 96.2%; Pred. No. 5.3e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```

```

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKA-EAELDKKADLPKNKVADLEKEISNLEILLG 58
Db 67 LAKKQTELEKLLDLDPEGTQDELKKA-EAELDKKADLPKNKVADLEKEISNLEILLG 126
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 127 GADSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 170

RESULT 4
Q9L583 ID Q9L583 PRELIMINARY; PRT; 252 AA.
AC Q9L583;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255544; AAF68097.1; -.
FT HSSP; P04268; 11C2.
FT NON_TER 1
FT NON_TER 252
SQ SEQUENCE 252 AA; 27260 MW; 82DE13741F369CA2 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 252;
Best Local Similarity 96.2%; Pred. No. 5.4e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKA-EAELDKKADLPKNKVADLEKEISNLEILLG 58
Db 73 LAKKQTELEKLLDLDPEGTQDELKKA-EAELDKKADLPKNKVADLEKEISNLEILLG 132
Qy 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 102
Db 133 GADSEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEE 176

RESULT 5
Q8KQK3 ID Q8KQK3 PRELIMINARY; PRT; 360 AA.
AC Q8KQK3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=259/98;
RX MEDLINE=22170754; PubMed=12183557;

```

```

RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.N., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082389; AAL92494.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 360
SQ SEQUENCE 360 AA; 39575 MW; 0C09A791547A47EC CRC64;

Query Match 92.7%; Score 468; DB 2; Length 360;
Best Local Similarity 96.2%; Pred. No. 7.6e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKQTQDELDKA-EAELDKKADLPNKVADLKEISNLEILLG 58
DB 232 LAKKQTELEKLLDLDPEKQTQDELDKAEEAELDKKADLPNKVADLKEISNLEILLG 291
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102
DB 292 GADSEDDTAALQNKLTAKKAELEKTKQELDAALNELGPDGDEE 335

RESULT 6
Q9LAX7 PRELIMINARY; PRT; 429 AA.
AC Q9LAX7;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071818; AAF27714.1; -.
DR InterPro; IPR009053; Prefoldin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 429
SQ SEQUENCE 429 AA; 47048 MW; BC1D74BBA54DA9D6 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 429;
Best Local Similarity 96.2%; Pred. No. 9e-24;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKQTQDELDKA-EAELDKKADLPNKVADLKEISNLEILLG 58
DB 254 LAKKQTELEKLLDLDPEKQTQDELDKAEEAELDKKADLPNKVADLKEISNLEILLG 313
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102
DB 314 GADSEDDTAALQNKLTAKKAELEKTKQELDAALNELGPDGDEE 357

RESULT 7
Q9LAX9 PRELIMINARY; PRT; 526 AA.
AC Q9LAX9;

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF3296;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071816; AAF27712.1; -.
DR HSSP; P04268; IIC2.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 526
SQ SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AE CRC64;

Query Match 92.7%; Score 468; DB 2; Length 526;
Best Local Similarity 96.2%; Pred. No. 1.1e-23;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKQTQDELDKA-EAELDKKADLPNKVADLKEISNLEILLG 58
DB 346 LAKKQTELEKLLDLDPEKQTQDELDKAEEAELDKKADLPNKVADLKEISNLEILLG 405
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102
DB 406 GADSEDDTAALQNKLTAKKAELEKTKQELDAALNELGPDGDEE 449

RESULT 8
Q8VQ55 PRELIMINARY; PRT; 608 AA.
AC Q8VQ55;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNH1156;
RA Lee K.J., Bae S.M., Chung K.S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF460993; AAL67804.1; -.
DR HSSP; P06653; 1HCX.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
FT NON_TER 608
FT NON_TER 608
SQ SEQUENCE 608 AA; 67918 MW; 15F71BD62E297526 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 608;
Best Local Similarity 96.2%; Pred. No. 1.3e-23;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEKQTQDELDKA-EAELDKKADLPNKVADLKEISNLEILLG 58
DB 222 LAKKQTELEKLLDLDPEKQTQDELDKAEEAELDKKADLPNKVADLKEISNLEILLG 281
QY 59 GADSEDDTAALPNKLTAKKAELEKTKQELDAALNELGPDGDEE 102

```

```
Db 282 GADSEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 325
RESULT 9
Q97T39 PRELIMINARY; PRT; 744 AA.
AC Q97T39;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A.
GN OrderedLocuNames=SP0117;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Mayhew L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
DR EMBL: AE007328; AAK74303.1; -.
DR PIR: F95013; F95013.
DR HSP: P06653; IHXC.
DR TIGR: SP0117; -.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR002345; Lipocalin.
DR Pfam: PF01473; CW binding 1; 10.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_2.
KW Complete proteome.
SQ SEQUENCE 744 AA; 82764 MW; 20E5E8E7911EFD5 CRC64;

Query Match 92.7%; Score 468; DB 2; Length 744;
Best Local Similarity 96.2%; Pred. No. 1.6e-23;
Matches 100; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 58
Db 346 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 405

Qy 59 GADSEDDTAALPNKATKKALEKTQKELDAALNELGPDGDEE 102
Db 406 GADSEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 449

RESULT 10
Q9L579 PRELIMINARY; PRT; 231 AA.
AC Q9L579;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP255547; AAF68100.1; -.
DR HSP: P04268; IIC2.
FT NON_TER 1 231
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26038 MW; BB87E1A4C25FA669 CRC64;

Query Match 91.7%; Score 463; DB 2; Length 241;
Best Local Similarity 95.2%; Pred. No. 1.1e-23;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 58
Db 77 LAKKQTELEKLLDNLDPGKTQDELKKEA-EAELDKKADLPKNKADLEKEINLEILLG 136

Qy 59 GADSEDDTAALPNKATKKALEKTQKELDAALNELGPDGDEE 102
Db 137 GADPEDDTAALQNKATKKALEKTQKELDAALNELGPDGDEE 180
```

```

RESULT 12
Q9L5B7
ID Q9L5B7 PRELIMINARY; PRT; 249 AA.
AC Q9L5B7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PcpA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253405; AAF67353.1; -.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 27271 MW; B4106707EF108A0B CRC64;

Query Match 91.7%; Score 463; DB 2; Length 249;
Best Local Similarity 95.2%; Pred. No. 1.1e-23;
Matches 99; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKPVADLKEISNLEILLG 58
DB 103 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPKPVADLKEISNLEILLG 162

QY 59 GADSEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 102
DB 163 GADPEDDTAALQNKLTATKKALEKTKQKELDAALNELGPDGDEE 206

RESULT 13
Q9LAX8
ID Q9LAX8 PRELIMINARY; PRT; 502 AA.
AC Q9LAX8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PcpA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8090;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PcpA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071817; AAF27713.1; -.
DR HSSP; O15813; 1D7M.
DR InterPro; IPR011047; Quin_abc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.

FT NON_TER 502
SQ SEQUENCE 502 AA; 55018 MW; 4E73D477CAE79B40 CRC64;

Query Match 90.9%; Score 459; DB 2; Length 502;
Best Local Similarity 94.2%; Pred. No. 4.2e-23;
Matches 98; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKPVADLKEISNLEILLG 58
DB 347 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPKPVADLKEISNLEILLG 406

QY 59 GADSEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 102
DB 407 GADPEDDTAALQNKLTATKKALEKTKQKELDAALNELGPDGDEE 450

RESULT 14
Q9L585
ID Q9L585 PRELIMINARY; PRT; 249 AA.
AC Q9L585;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PcpA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Packlam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RN J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255542; AAF68095.1; -.
DR HSSP; P04268; 1IC2.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 27050 MW; DF4D2ED9265986FA CRC64;

Query Match 90.5%; Score 457; DB 2; Length 249;
Best Local Similarity 94.2%; Pred. No. 2.9e-23;
Matches 98; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLD-LDPEGKTQDELKKEA-EAELDKKADLPKPVADLKEISNLEILLG 58
DB 68 LAKKQTELEKLLDLDPEGKTQDELKKEA-EAELDKKADLPKPVADLKEISNLEILLG 127

QY 59 GADSEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 102
DB 128 GADPEDDTAALQNKLTATKKALEKTKQKELDAALNELGPDGDEE 171

RESULT 15
Q9L590
ID Q9L590 PRELIMINARY; PRT; 256 AA.
AC Q9L590;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PcpA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

```

|    |  |
|----|--|
| OC | Streptococcus.   |
| OX | NCBI_TaxID=1313;   |
| RN | [1]  |
| RP | SEQUENCE FROM N.A.   |
| RC | STRAIN=SP193;  |
| RX | MEDLINE=20472698; PubMed=11015380;                               |
| RA | Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;          |
| RT | "Pneumococcal pspa sequence types of prevalent multiresistant    |
| RT | pneumococcal strains in the United States and of internationally |
| RT | disseminated clones.";   |
| RL | J. Clin. Microbiol. 38:3663-3669(2000).                          |

Search completed: June 21, 2005, 10:22:13  
Job time : 63.1776 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 59.6735 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-19

Perfect score: 397

Sequence: 1 LDKEAGEALDKADGLPNK.....TQKELDAALNELGPDGDEE 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

- Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseq19808:\*

2: Geneseq19908:\*

3: Geneseq20008:\*

4: Geneseq20018:\*

5: Geneseq20028:\*

6: Geneseq20038a:\*

7: Geneseq20038b:\*

8: Geneseq20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID         | Description        |
|------------|-------|-------|--------|------------|--------------------|
| 1          | 397   | 100.0 | 197    | 7 ABW02598 | Abw02598 Ac122c pn |
| 2          | 397   | 100.0 | 8991   | 6 ABU08487 | Abu08487 S. pneumo |
| 3          | 381.5 | 96.1  | 196    | 2 AAW14564 | Aaw14564 Streptoco |
| 4          | 368   | 92.7  | 233    | 2 AAW14572 | Aaw14572 Streptoco |
| 5          | 368   | 92.7  | 233    | 7 ABW02606 | Abw02606 Ef1019c p |
| 6          | 365   | 91.9  | 213    | 7 ABW02601 | Abw02601 Bg8090c p |
| 7          | 363   | 91.4  | 213    | 2 AAW14567 | Aaw14567 Streptoco |
| 8          | 362   | 91.2  | 416    | 8 ADK52498 | Adk52498 alpha hel |
| 9          | 362   | 91.2  | 526    | 8 ADK52497 | Adk52497 PspA mole |
| 10         | 362   | 91.2  | 744    | 6 ABU00449 | Abu00449 S. pneumo |
| 11         | 362   | 91.2  | 744    | 8 ADM2054  | Adm2054 S. pneumon |
| 12         | 362   | 90.2  | 745    | 3 AAY81652 | Aay81652 Streptoco |
| 13         | 358   | 90.2  | 641    | 2 AAW61217 | Aaw61217 Streptoco |
| 14         | 358   | 90.2  | 641    | 5 ABP54636 | Abp54636 S. pneumo |
| 15         | 358   | 90.2  | 641    | 7 ADC45241 | Adc45241 S. pneumo |
| 16         | 238   | 59.9  | 211    | 7 ABW02621 | Abw02621 Bg11703c  |
| 17         | 238   | 59.9  | 238    | 2 AAW14587 | Aaw14587 Streptoco |
| 18         | 235   | 59.2  | 212    | 2 AAW14588 | Aaw14588 Streptoco |
| 19         | 235   | 59.2  | 212    | 7 ABW02622 | Abw02622 Bg7817c p |
| 20         | 233   | 58.7  | 232    | 7 ABW02624 | Abw02624 Ef5668c p |
| 21         | 233   | 58.7  | 233    | 2 AAW14590 | Aaw14590 Streptoco |
| 22         | 233   | 58.7  | 275    | 8 ADOS2055 | Ados2055 S. pneumo |
| 23         | 233   | 58.7  | 369    | 8 ADK52496 | Adk52496 alpha hel |
| 24         | 233   | 58.7  | 458    | 2 AAW14592 | Aaw14592 Streptoco |
| 25         | 233   | 58.7  | 458    | 7 ABW02626 | Abw02626 Ef5668 pn |

RESULT 1

ABW02598

ID ABW02598 standard; protein; 197 AA.

XX AC ABW02598;

XX DT 12-FEB-2004 (first entry)

XX DE Ac122c pneumococcal surface protein A (PspA) central region.

XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;

XX KW immunological; gene therapy; immunostimulant.

XX OS Unidentified.

XX PN US6592876-B1.

XX PD 15-JUL-2003.

XX PF 15-SEP-1995; 95US-00529055.

XX PR 20-APR-1993; 93US-00048896.

XX PR 06-JUN-1995; 95US-00465746.

XX (UABR-) UAB RES FOUND.

XX PI Briles DE, Medaniel LS, Swiatlo E, Yother J, Brooks-Walter A;

XX DR WPI; 2003-862841/80.

XX PT Immunological composition for obtaining expression products used for detecting the presence of Streptococcus pneumoniae or its strain, comprises at least two different full length isolated gene encoding pneumococcal surface protein A.

XX Example 6; SEQ ID NO 44; 121pp; English.

XX The present invention relates to an immunological composition comprising at least 2 different full length isolated genes encoding pneumococcal surface protein A (PspA) from different groups based on restriction fragment polymorphism analysis. The invention is useful for obtaining expression products by recombinant techniques to detect, determine, isolate or diagnose the presence of Streptococcus pneumoniae or its strain. The expression product is useful for preparing antigenic, immunological or vaccine compositions, for eliciting antibodies, an immunological response (other than or additional to antibodies), or a protective response (including antibody or other immunological response by administering compositions to a host). The invention is also useful as

CC vaccines and in gene therapy. The present sequence is Acl22c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention

XX Sequence 197 AA;

Query Match 100.0%; Score 397; DB 7; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-33;  
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60  
 |||||  
 Db 46 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 105  
 |||||

QY 61 TOKELDAALNELGPDGDEEE 80  
 |||||  
 Db 106 TOKELDAALNELGPDGDEEE 125  
 |||||

RESULT 2  
 ID ABU08487  
 AC ABU08487 standard; protein; 8991 AA.

XX ABU08487;

XX 24-JUN-2003 (first entry)

XX S. pneumoniae pneumococcal surface protein A (PspA) protein.

XX Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KW alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KW antibacterial.

XX Streptococcus pneumoniae.

XX Key Location/Qualifiers

FT Misc-difference 1..8991  
 FT /note= "All Xaa residues within this sequence are  
 FT unknown"

XX US6500613-B1.

XX 31-DEC-2002.

XX 16-SEP-1996; 96US-00714741.

XX 15-SEP-1995; 95US-00529055.

XX (UYAL-) UNIV ALABAMA.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 2003-361534/34.

XX Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.

XX Disclosure; Col 145-188; 186pp; English.

XX The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents S. pneumoniae  
 CC PspA protein

XX Sequence 8991 AA;

Query Match 100.0%; Score 397; DB 6; Length 8991;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-31;  
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60  
 |||||  
 Db 4108 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 4167  
 |||||  
 QY 61 TOKELDAALNELGPDGDEEE 80  
 |||||  
 Db 4168 TOKELDAALNELGPDGDEEE 4187  
 |||||

RESULT 3

AAW14564

ID AAW14564 standard; protein; 196 AA.

XX AAW14564;

XX 17-OCT-2003 (revised)

XX 28-OCT-1997 (first entry)

XX Streptococcus pneumoniae PspA central region.

XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.

XX Streptococcus pneumoniae; strain Acl22.

XX WO9709994-A1.

XX 20-MAR-1997.

XX 16-SEP-1996; 96WO-US014819.

XX 15-SEP-1995; 95US-00529055.

XX (UABR-) UAB RES FOUND.

XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Acl22.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 196 AA;

Query Match 96.1%; Score 381.5; DB 2; Length 196;  
 Best Local Similarity 98.8%; Pred. No. 2.6e-31;

Matches 79; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60  
 |||||

Db 46 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 105  
 |||||



```

QY 61 TQKELDAALNELGPDGDEEE 80
Db 106 T-KELDAALNELGPDGDEEE 124

RESULT 4
AAW14572
ID AAW14572 standard; protein; 233 AA.
XX
AC AAW14572;
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
DE Streptococcus pneumoniae PspA central region.
XX
KW PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
OS Streptococcus pneumoniae; strain EF3296.
XX
FH Key Location/Qualifiers
FT Misc-difference 129 /note= "unidentified amino acid"
FT Misc-difference 131 /note= "unidentified amino acid"
FT
XX
PN WO9709994-A1.
XX
PD 20-MAR-1997.
XX
PF 16-SEP-1996; 96WO-US014819.
XX
PR 15-SEP-1995; 95US-00529055.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
PI Hollingshead S, Tart R, Brooks-Walter A;
XX
DR WPI; 1997-202002/18.
XX
ST Streptococcus pneumoniae surface protein PspC and truncated PspA - used
PT in vaccines for protecting animals against S.pneumoniae infection.
XX
PS Example 6; Fig 13; 296pp; English.
XX
CC This sequence shows the central portion, including the C-terminus of the
CC alpha-helix region and some of the proline-rich region, of pneumococcal
CC surface protein A (PspA) of Streptococcus pneumoniae strain EF3296.
CC Comparison of the N-terminal and central regions (AAW14533-57 and
CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can
CC be used to divide the strains into several families based on sequence
CC homologies. PspA polypeptides, or fragments of them, can be used in
CC vaccines to protect animals against S. pneumoniae infection and hence for
CC the prevention of diseases such as otitis media, meningitis, bacteraemia
CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical
CC region and the immediate 5' tip of the coding sequence are likely to be
CC the critical sequences for predicting PspA cross-reactions and vaccine
CC composition. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 233 AA;

Query Match 92.7%; Score 368; DB 2; Length 233;
Best Local Similarity 93.8%; Pred. No. 7.9e-30;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 77 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 136
QY 61 TQKELDAALNELGPDGDEEE 80
Db 137 TQKELDAALNELGPDGDEEE 156

RESULT 5
ABW02606
ID ABW02606 standard; protein; 233 AA.
XX
AC ABW02606;
XX
DT 12-FEB-2004 (first entry)
XX
DE Ef1019c pneumococcal surface protein A (PspA) central region.
XX
KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .233 /note= "Xaa = Unknown amino acid"
FT
XX
PN US6592876-B1.
XX
PD 15-JUL-2003.
XX
PF 15-SEP-1995; 95US-00529055.
XX
PR 20-APR-1993; 93US-00048896.
XX
PR 06-JUN-1995; 95US-00465746.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
DR WPI; 2003-862841/80.
XX
ST Immunological composition for obtaining expression products used for
PT detecting the presence of Streptococcus pneumoniae or its strain,
PT comprises at least two different full length isolated gene encoding
PT pneumococcal surface protein A.
XX
PS Example 6; SEQ ID NO 52; 121pp; English.
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspAs) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antigenic,
CC immunological or vaccine compositions, for eliciting antibodies, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Ef1019c
CC pneumococcal surface protein A (PspA) central region. This sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 233 AA;

Query Match 92.7%; Score 368; DB 7; Length 233;
Best Local Similarity 93.8%; Pred. No. 7.9e-30;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 77 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 136
QY 61 TQKELDAALNELGPDGDEEE 80
Db 137 TQKELDAALNELGPDGDEEE 156

```

|                       |   |
|-----------------------|---|
| RESULT 6              |   |
| ABW02601              |   |
| ID                    | ABW02601 standard; protein; 213 AA.                                       |
| XX                    |   |
| AC                    | ABW02601;   |
| XX                    |   |
| DT                    | 12-FEB-2004 (first entry)   |
| XX                    |   |
| DE                    | Bg8090c pneumococcal surface protein A (PspA) central region.             |
| XX                    |   |
| KW                    | Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;      |
| XX                    | immunological; gene therapy; immunostimulant.                             |
| XX                    |   |
| OS                    | Unidentified.   |
| XX                    |   |
| Key                   | Location/Qualifiers   |
| FT                    | Misc-difference 2   |
| FT                    | /label= Unknown   |
| XX                    |   |
| PN                    | US6592876-B1.   |
| XX                    |   |
| PD                    | 15-JUL-2003.  |
| XX                    |   |
| PF                    | 15-SEP-1995; 95US-00529055.   |
| XX                    |   |
| PR                    | 20-APR-1993; 93US-00048896.   |
| PR                    | 06-JUN-1995; 95US-00465746.   |
| XX                    |   |
| PA                    | (UABR-) UAB RES FOUND.  |
| XX                    |   |
| PI                    | Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;             |
| XX                    |   |
| DR                    | WPI; 2003-862841/80.  |
| XX                    |   |
| PT                    | Immunological composition for obtaining expression products used for      |
| PT                    | detecting the presence of Streptococcus pneumoniae or its strain,         |
| PT                    | comprises at least two different full length isolated gene encoding       |
| PT                    | pneumococcal surface protein A.   |
| XX                    |   |
| PS                    | Example 6; SEQ ID NO 47; 121pp; English.                                  |
| XX                    |   |
| CC                    | The present invention relates to an immunological composition comprising  |
| CC                    | at least 2 different full length isolated genes encoding pneumococcal     |
| CC                    | surface protein A (pspas) from different groups based on restriction      |
| CC                    | fragment polymorphism analysis. The invention is useful for obtaining     |
| CC                    | expression products by recombinant techniques to detect, determine,       |
| CC                    | isolate or diagnose the presence of Streptococcus pneumoniae or its       |
| CC                    | strain. The expression product is useful for preparing antigenic, an      |
| CC                    | immunological or vaccine compositions, for eliciting antibodies, an       |
| CC                    | immunological response (other than or additional to antibodies) or a      |
| CC                    | protective response (including antibody or other immunological response   |
| CC                    | by administering compositions to a host). The invention is also useful as |
| CC                    | vaccines and in gene therapy. The present sequence is Bg8090c             |
| CC                    | pneumococcal surface protein A (PspA) central region. This sequence is    |
| CC                    | used in the exemplification of the invention                              |
| XX                    |   |
| SQL                   | Sequence 213 AA;  |
|                       |   |
| Query Match           | 91.9%; Score 365; DB 7; Length 213;                                       |
| Best Local Similarity | 92.5%; Pred. No. 1.4e-29;   |
| Matches               | 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;                       |
|                       |   |
| QY                    | 1 LDKAAGBAELDKKADGLPNKVSQDLEKISINLEILLGGADSEDDTAALPNKLTATKAELEK 60        |
|                       |   |
| DB                    | 83 LDKAABAELDKKADGLPNKVAULEKISINLEILLGGADPEDDTAALPNKLTATKABFEK 142        |
|                       |   |
| QY                    | 61 TKELDAALNELGPDGDEEE 80   |
|                       |   |
| DB                    | 143 TPKELDAALNELGPDGDEEE 162  |
|                       |   |
| RESULT 7              |   |

```

XX 20-MAY-2004 (first entry)
XX alpha helical region PspA molecule from the EF3296 strain.
XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
XX hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
XX Hodgkin's disease.
XX Streptococcus pneumoniae.
XX WO2004016231-A2.
XX 26-FEB-2004.
XX 17-FEB-2003; 2003WO-US008199.
XX 15-MAR-2002; 2002US-0365351P.
XX (UABR-) UAB RES FOUND.
XX Briles DE;
XX WPI; 2004-192068/18.
XX Treating Streptococcus pneumoniae infection in a subject lacking a
XX functional spleen comprises administering an antibody that recognizes
XX pneumococcal surface protein A (PspA) or its binding portion.
XX Claim 17; SEQ ID NO 4; 41pp; English.
XX The present invention relates to treating Streptococcus pneumoniae
XX infection in a subject lacking a functional spleen comprises
XX administering an antibody that recognizes pneumococcal surface protein A
XX (PspA) or its binding portion. The method is useful for treating or
XX preventing Streptococcus pneumoniae infection in a subject lacking a
XX functional spleen. The disease-associated injury is especially due to
XX hemolytic anemia disease, leukemia or lymphoma, especially sickle cell
XX anemia or Hodgkin's disease. The present sequence represents the alpha
XX helical region PspA molecule from the EF3296 strain of Streptococcus
XX pneumoniae.
XX Sequence 416 AA;
XX Query Match 91.2%; Score 362; DB 8; Length 416;
XX Best Local Similarity 93.8%; Pred. No. 6.6e-29;
XX Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX 1 LDKEAGEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
XX 265 LDKEAEAEALDKKADGLQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 324
XX 61 TOKELDAALNELGPDGDEE 80
XX 325 TOKELDAALNELGPDGDEE 344
XX RESULT 9
XX ADK52497
XX ID ADK52497 standard; protein; 526 AA.
XX AC ADK52497;
XX DT 20-MAY-2004 (first entry)
XX DE PspA molecule from the EF3296 strain of Streptococcus pneumoniae.
XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;
XX hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;
XX Hodgkin's disease.
XX Streptococcus pneumoniae.
XX Query Match 91.2%; Score 362; DB 8; Length 526;
XX Best Local Similarity 93.8%; Pred. No. 8.9e-29;
XX Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX 1 LDKEAGEALDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
XX 370 LDKEAEAEALDKKADGLQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTATKKALEK 429
XX 61 TOKELDAALNELGPDGDEE 80
XX 430 TOKELDAALNELGPDGDEE 449
XX RESULT 10
XX ABU00449
XX ID ABU00449 standard; protein; 744 AA.
XX AC ABU00449;
XX DT 23-OCT-2003 (revised)
XX DT 11-FEB-2003 (first entry)
XX DE S. pneumoniae type 4 strain protein from coding region #16.
XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
XX antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
XX gene therapy; vaccine.
XX Streptococcus pneumoniae; type 4 strain.
XX WO200277021-A2.
XX 03-OCT-2002.
XX 27-MAR-2002; 2002WO-IB002163.
XX 27-MAR-2001; 2001GB-00007658.
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.

```

```

PI Masignani V, Tettelin H, Fraser C;
XX WPI; 2003-040579/03.
DR N-PSDB; ABX05728.
DR
XX
XX
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX
XX Claim 1; SEQ ID NO 32; 56pp; English.
PS
XX
XX The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC ADS56454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the proteins has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to streptococcus
CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is one of the 2469 proteins
CC expressed by the identified coding regions from the genomic sequence.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
XX standardise OS field)
XX
SQ Sequence 744 AA;

Query Match          91.2%; Score 362; DB 6; Length 744;
Best Local Similarity 93.8%; Pred. No. 1.4e-28;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 LDKEAGEAELDKKADGLPNKVSDLEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 LDKEAGEAELDKKADGLPNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 429

Oy 61 TQKELDAALNELGPDGDEEE 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 TQKELDAALNELGPDGDEEE 449

RESULT 12
AAY81652
ID AAY81652 standard; protein; 745 AA.
XX
XX AAY81652;
DT 24-MAY-2000 (first entry)
XX
DE Streptococcus pneumoniae protein sequence ID301.
XX
XX Streptococcus pneumoniae; vaccine; screening; protein antigen;
XX antibacterial; antiinflammatory; meningitis; infection; diagnosis;
XX pneumococcal disease.
XX
XX Streptococcus pneumoniae.
XX
XX WO200006737-A2.
XX
XX 10-FEB-2000.
XX
XX 27-JUL-1999; 99WO-GB002451.
XX
XX 27-JUL-1998; 98GB-00016337.
XX 19-MAR-1999; 99US-0125164P.
XX
XX (MICR-) MICROBIAL TECHNICS LTD.
XX
XX Gilbert CFG, Hansbro PM;
XX

```

DR WPI; 2000-195300/17.  
XX New Streptococcal protein, useful as a vaccine, for diagnosis of  
PT pneumococcal diseases and for screening agents capable of antagonizing or  
PT inhibiting expression of the protein.  
XX  
XX Claim 2; Page 95; 108pp; English.  
XX  
XX AAY81501 to AAY81679 represent specifically claimed protein sequences  
CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent  
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.  
CC The sequences have antibacterial and antiinflammatory properties. The  
CC protein sequences, and fragments of them, are useful as immunogens and/or  
CC antigens. The nucleotide sequences can be used in vaccines and in  
CC diagnostic assays. The proteins and nucleotides can be useful for the  
CC detection and diagnosis of S. pneumoniae. The protein sequences are also  
CC useful for screening an agent capable of antagonising, inhibiting or  
CC interfering with the function or expression of the proteins in which the  
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection  
CC and meningitis. AAA05591 to AAA05614 represent primers used in the  
CC exemplification of the present invention  
XX  
XX Sequence 745 AA;  
SQ  
Query Match 91.2%; Score 362; DB 3; Length 745;  
Best Local Similarity 93.8%; Pred. No. 1.4e-28;  
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60  
DB 370 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 429  
QY 61 TQKELDAALNELGPDGDEEE 80  
DB 430 TQKELDAALNELGPDGDEEE 449  
RESULT 13  
AAW61217  
ID AAW61217 standard; protein; 641 AA.  
XX  
XX AAW61217;  
AC  
XX 02-OCT-1998 (first entry)  
DE  
XX Streptococcus pneumoniae SP0092 protein.  
DE  
XX Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;  
KW Streptococcus pneumoniae; otitis media; meningitis.  
KW detection; pneumonia; otitis media; meningitis.  
XX  
XX Streptococcus pneumoniae.  
XX  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 306  
FT /label= unknown  
FT /note= "encoded by NCT"  
FT  
XX  
XX WO9818930-A2.  
XX  
XX 07-MAY-1998.  
XX  
XX 30-OCT-1997; 97WO-US019422.  
XX  
XX 31-OCT-1996; 96US-0029960P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Kunsch CA, Choi GH, Johnson LS, Hromockyj A;  
XX  
XX WPI; 1998-272224/24.  
XX N-PSDB; AAV27403.  
XX  
XX Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae

PT - or their epitope-containing fragments, useful in protective or  
therapeutic vaccines, and for diagnosis.  
XX  
XX Claim 11; Page 82; 118pp; English.  
XX  
XX The present sequence represents a protein from Streptococcus pneumoniae.  
CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein  
CC can be useful in vaccines for inducing protective antibodies against  
CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.  
CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
CC are used to detect Streptococcus infection (by usual hybridisation or  
CC amplification methods), also for isolating Streptococcus genes or their  
CC allelic variants. The protein can be used similarly to detect specific  
CC antibodies in standard immunoassays, especially for diagnosing or  
CC monitoring infections. Antibodies which bind the protein are used to  
CC detect corresponding antigens, to purify the protein and for passive  
CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
CC e.g. by injection, orally or through the skin, typically at 0.01-1000  
CC (especially 10-300) mu g/ml per dose  
XX  
XX Sequence 641 AA;  
SQ  
Query Match 90.2%; Score 358; DB 2; Length 641;  
Best Local Similarity 92.5%; Pred. No. 2.9e-28;  
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60  
DB 267 LDKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 326  
QY 61 TQKELDAALNELGPDGDEEE 80  
DB 327 TQKELDAALNELGPDGDEEE 346  
RESULT 14  
ABP54636  
ID ABP54636 standard; protein; 641 AA.  
XX  
XX ABP54636;  
AC  
XX 04-SEP-2002 (first entry)  
DT  
XX S. pneumoniae SP092 protein sequence SEQ ID NO:160.  
DE  
XX Streptococcus pneumoniae; epitope; vaccine; antigenic protein;  
KW antibacterial; Streptococcal infection; detection.  
KW  
XX Streptococcus pneumoniae.  
XX  
XX US2002061545-A1.  
XX  
XX 23-MAY-2002.  
PD  
XX 22-JAN-2001; 2001US-00765272.  
PF  
XX 30-OCT-1997; 97US-00961083.  
PR  
XX (CHOI/) CHOI G H.  
PA (KUNS/) KUNSCH C A.  
PA (BARA/) BARASH S C.  
PA (DILL/) DILLON P J.  
PA (DOUG/) DOUGHERTY B.  
PA (FANN/) FANNON M R.  
PA (ROSE/) ROSEN C A.  
XX  
XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;  
PI Rosen CA;  
XX  
XX WPI; 2002-479261/51.  
DR N-PSDB; ABQ84871.  
DR  
XX New Streptococcus pneumoniae antigens, useful for detecting Streptococcus

PT and for preventing or attenuating disease caused by Streptococcus  
 XX infection.

PS Claim 11; Page 43; 70pp; English.

XX ABQ84792 to ABQ84904 represents nucleic acids which encode the  
 CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669. The S.  
 CC pneumoniae antigens have antibacterial activity and can be used in  
 CC vaccines. The S. pneumoniae antigens can also be used to prevent or  
 CC attenuate a Streptococcal infection in an animal. The polynucleotides  
 CC encoding the S. pneumoniae antigens can be used to detect Streptococcus  
 CC nucleic acids. ABQ84905 to ABQ85130 represent primers used in the cloning  
 CC of S. pneumoniae ORFs (open reading frames) which are used in an example  
 CC from the present invention

XX Sequence 641 AA;

Query Match 90.2%; Score 358; DB 5; Length 641;  
 Best Local Similarity 92.5%; Pred. No. 2.9e-28;  
 Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60  
 DB 267 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 326

QY 61 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 80

DB 327 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 326

RESULT 15

ID ADC45241  
 ID ADC45241 standard; protein; 641 AA.

AC ADC45241;

XX 18-DEC-2003 (first entry)

XX S. pneumoniae antigenic protein SP092.

XX Antigen; bacterial infection; vaccine; pneumonia; antibacterial.

XX Streptococcus pneumoniae.

XX US6573082-B1.

XX 03-JUN-2003.

XX 28-MAR-2000; 2000US-00536784.

XX 31-OCT-1996; 96US-0029960P.

XX 30-OCT-1997; 97US-00961083.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;  
 PI Roen CA;

XX WPI; 2003-764574/72.

XX N-PSDB; ADC45240.

XX Novel polynucleotide encoding Streptococcus pneumoniae polypeptides  
 PT useful for producing vaccines for prevention or attenuation of infection  
 PT by Streptococcus pneumoniae.

XX Example 1; SEQ ID NO 160; 58pp; English.

XX The invention relates to an isolated polynucleotide consisting of a  
 CC Streptococcus pneumoniae nucleic acid (appearing as ADC45122 and encoding  
 CC SP028) one of 113 disclosed nucleic acids encoding 113 S. pneumoniae  
 CC antigens. Also included are making a recombinant vector by inserting the  
 CC nucleic acid into a vector, an isolated polynucleotide consisting of at  
 CC least 50 or 100 contiguous nucleotides of the SP028 nucleic acid, and a

CC recombinant host cell comprising the SP028 polynucleotide. The nucleic  
 CC acids are useful as DNA vaccine against Streptococcus pneumoniae  
 CC infection (e.g. pneumonia). Nucleic acids derived from the S. pneumoniae  
 CC antigen nucleic acids are useful as probes for use in diagnostic methods  
 CC for detecting S. pneumoniae gene expression. The present sequence  
 XX represents an S. pneumoniae antigenic protein.

SQ Sequence 641 AA;

Query Match 90.2%; Score 358; DB 7; Length 641;  
 Best Local Similarity 92.5%; Pred. No. 2.9e-28;  
 Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60  
 DB 267 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 326

QY 61 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 80

DB 327 LDKKAGEAEELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 326

Search completed: June 21, 2005, 10:10:17  
 Job time : 59.6735 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:55:14 ; Search time 14.9388 Seconds  
(without alignments)  
399.760 Million cell updates/sec

Title: US-10-674-755-19  
Perfect score: 397  
Sequence: 1 LDKAGEALDKKADGLPNK.....TQKELDAALNELGPDGDEEE 80

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCUTS COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 397   | 100.0       | 80     | 2  | US-08-710-749-18  |
| 2          | 397   | 100.0       | 80     | 4  | US-09-147-875A-19 |
| 3          | 397   | 100.0       | 197    | 4  | US-08-529-055-44  |
| 4          | 397   | 100.0       | 8991   | 4  | US-08-714-741-32  |
| 5          | 378   | 95.2        | 104    | 2  | US-08-710-749-19  |
| 6          | 378   | 95.2        | 104    | 4  | US-09-147-875A-20 |
| 7          | 375   | 94.5        | 104    | 4  | US-09-147-875A-21 |
| 8          | 369.5 | 93.1        | 102    | 2  | US-08-710-749-21  |
| 9          | 369.5 | 93.1        | 102    | 4  | US-08-710-749-18  |
| 10         | 369   | 92.9        | 104    | 2  | US-08-710-749-20  |
| 11         | 368   | 92.7        | 233    | 4  | US-08-529-055-52  |
| 12         | 365   | 91.9        | 213    | 4  | US-08-529-055-47  |
| 13         | 358   | 90.2        | 641    | 3  | US-08-961-083-160 |
| 14         | 358   | 90.2        | 641    | 4  | US-09-536-784-160 |
| 15         | 238   | 59.3        | 108    | 2  | US-08-710-749-22  |
| 16         | 238   | 59.9        | 108    | 2  | US-08-710-749-23  |
| 17         | 238   | 59.9        | 108    | 2  | US-08-710-749-26  |
| 18         | 238   | 59.9        | 108    | 4  | US-09-147-875A-23 |
| 19         | 238   | 59.9        | 211    | 4  | US-08-529-055-67  |
| 20         | 235   | 59.2        | 108    | 4  | US-09-147-875A-24 |
| 21         | 235   | 59.2        | 212    | 4  | US-08-529-055-68  |
| 22         | 233   | 58.7        | 108    | 2  | US-08-710-749-24  |
| 23         | 233   | 58.7        | 108    | 4  | US-09-147-875A-25 |
| 24         | 233   | 58.7        | 232    | 4  | US-08-529-055-70  |
| 25         | 233   | 58.7        | 458    | 4  | US-08-529-055-73  |
| 26         | 225   | 56.7        | 106    | 4  | US-09-147-875A-22 |
| 27         | 186   | 46.9        | 108    | 2  | US-08-710-749-25  |

|    |       |      |     |   |                   |                   |
|----|-------|------|-----|---|-------------------|-------------------|
| 28 | 186   | 46.9 | 108 | 4 | US-09-147-875A-26 | Sequence 26, Appl |
| 29 | 186   | 46.9 | 185 | 4 | US-08-529-055-69  | Sequence 69, Appl |
| 30 | 149.5 | 37.7 | 119 | 2 | US-08-710-749-27  | Sequence 27, Appl |
| 31 | 149.5 | 37.7 | 119 | 4 | US-09-147-875A-27 | Sequence 27, Appl |
| 32 | 149.5 | 37.7 | 215 | 4 | US-08-529-055-43  | Sequence 43, Appl |
| 33 | 109.5 | 27.6 | 204 | 4 | US-08-529-055-51  | Sequence 51, Appl |
| 34 | 106.5 | 26.8 | 195 | 4 | US-08-529-055-71  | Sequence 71, Appl |
| 35 | 106   | 26.7 | 623 | 4 | US-08-714-741-47  | Sequence 47, Appl |
| 36 | 105.5 | 26.6 | 99  | 2 | US-08-710-749-10  | Sequence 10, Appl |
| 37 | 105.5 | 26.6 | 99  | 2 | US-08-710-749-11  | Sequence 11, Appl |
| 38 | 105.5 | 26.6 | 99  | 4 | US-09-147-875A-11 | Sequence 11, Appl |
| 39 | 105.5 | 26.6 | 198 | 4 | US-08-529-055-61  | Sequence 61, Appl |
| 40 | 105.5 | 26.6 | 288 | 3 | US-08-312-949-4   | Sequence 4, Appl  |
| 41 | 105.5 | 26.6 | 288 | 3 | US-08-446-201-4   | Sequence 4, Appl  |
| 42 | 105.5 | 26.6 | 289 | 1 | US-08-072-070-4   | Sequence 4, Appl  |
| 43 | 105.5 | 26.6 | 289 | 1 | US-08-469-434-4   | Sequence 4, Appl  |
| 44 | 105.5 | 26.6 | 289 | 1 | US-08-214-222-4   | Sequence 4, Appl  |
| 45 | 105.5 | 26.6 | 289 | 2 | US-08-467-852A-5  | Sequence 5, Appl  |

ALIGNMENTS

RESULT 1  
US-08-710-749-18  
; Sequence 18, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 80 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-18

Query Match 100.0%; Score 397; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.3e-37;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LDKAGEALDKKADGLPNKVDLKEISNLIILGGADSEDDTAALPNKLTAKKAELEK 60  
DB 1 LDKAGEALDKKADGLPNKVDLKEISNLIILGGADSEDDTAALPNKLTAKKAELEK 60





```
Query Match      100.0%; Score 397; DB 4; Length 8991;
Best Local Similarity 100.0%; Pred. No. 3.9e-34; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 60
    |||||
Db 4108 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 4167
    |||||

QY 61 TOKELDAALNELGPDGDEEE 80
    |||||
Db 4168 TOKELDAALNELGPDGDEEE 4187
    |||||

RESULT 5
US-08-710-749-19
; Sequence 19, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-19

Query Match      95.2%; Score 378; DB 2; Length 104;
Best Local Similarity 96.2%; Pred. No. 1.5e-34;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 60
    |||||
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 84
    |||||

QY 61 TOKELDAALNELGPDGDEEE 80
    |||||
Db 85 TOKELDAALNELGPDGDEEE 104
    |||||

RESULT 6
US-09-147-875A-20
; Sequence 20, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,875A
; FILING DATE: 1999-05-24
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-21

Query Match      94.5%; Score 375; DB 4; Length 104;
Best Local Similarity 95.0%; Pred. No. 3.2e-34;
Matches 76; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 60
    |||||
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 84
    |||||

QY 61 TOKELDAALNELGPDGDEEE 80
    |||||
Db 85 TOKELDAALNELGPDGDEEE 104
    |||||

RESULT 7
US-09-147-875A-21
; Sequence 21, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: Becker et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,875A
; FILING DATE: 1999-05-24
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-21

Query Match      95.2%; Score 378; DB 4; Length 104;
Best Local Similarity 96.2%; Pred. No. 1.5e-34;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 60
    |||||
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 84
    |||||

QY 61 TOKELDAALNELGPDGDEEE 80
    |||||
Db 85 TOKELDAALNELGPDGDEEE 104
    |||||

RESULT 8
US-08-710-749-21
; Sequence 21, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
```

```
GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-20

Query Match      95.2%; Score 378; DB 4; Length 104;
Best Local Similarity 96.2%; Pred. No. 1.5e-34;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 60
    |||||
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 84
    |||||

QY 61 TOKELDAALNELGPDGDEEE 80
    |||||
Db 85 TOKELDAALNELGPDGDEEE 104
    |||||

RESULT 7
US-09-147-875A-21
; Sequence 21, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-21

Query Match      94.5%; Score 375; DB 4; Length 104;
Best Local Similarity 95.0%; Pred. No. 3.2e-34;
Matches 76; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 60
    |||||
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKAELEK 84
    |||||

QY 61 TOKELDAALNELGPDGDEEE 80
    |||||
Db 85 TOKELDAALNELGPDGDEEE 104
    |||||

RESULT 8
US-08-710-749-21
; Sequence 21, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
```

```
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-21

Query Match          93.1%; Score 369.5; DB 2; Length 102;
Best Local Similarity 96.2%; Pred. No. 1.3e-33;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 LDKEAGEAEELDKKADGPNKVKSDLEKEISNLEILGGADSEDDTAALPNKLTAKKAELEK 60
Db 24 LDKEA-EAEELDKKADGPNKVKADLEKEISNLEILGGADSEDDTAALPNKLTAKKAELEK 82

QY 61 TQKELDAALNELGPDGDEEE 80
Db 83 TQKELDAALNELGPDGDEEE 102

RESULT 9
US-09-147-875A-18
; Sequence 18, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-18

Query Match          93.1%; Score 369.5; DB 4; Length 102;
Best Local Similarity 96.2%; Pred. No. 1.3e-33;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 LDKEAGEAEELDKKADGPNKVKSDLEKEISNLEILGGADSEDDTAALPNKLTAKKAELEK 60
Db 24 LDKEA-EAEELDKKADGPNKVKADLEKEISNLEILGGADSEDDTAALPNKLTAKKAELEK 82

QY 61 TQKELDAALNELGPDGDEEE 80
Db 83 TQKELDAALNELGPDGDEEE 102

RESULT 10
```

```
US-08-710-749-20
; Sequence 20, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-20

Query Match          92.9%; Score 369; DB 2; Length 104;
Best Local Similarity 93.8%; Pred. No. 1.5e-33;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKEAGEAEELDKKADGPNKVKSDLEKEISNLEILGGADSEDDTAALPNKLTAKKAELEK 60
Db 25 LDKEA-EAEELDKKADGPNKVKADLEKEISNLEILGGADSEDDTAALPNKLTAKKAELEK 84

QY 61 TQKELDAALNELGPDGDEEE 80
Db 85 TQKELDAALNELGPDGDEEE 104

RESULT 11
US-08-529-055-52
; Sequence 52, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
```

```

; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-52

Query Match          92.7%; Score 368; DB 4; Length 233;
Best Local Similarity 93.8%; Pred. No. 5.4e-33;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 LDKEAGEAELDKKADGLPNKVSDLEKEISNLEILLGGADSEDDTAALPNKLATKKAELEK 60
        |||||
DB      77 LDKEAGEAELDKKADGLPNKVADLEKEISNLEILLGGADSEDDTAALPNKLAKXKALEK 136
        |||||

QY      61 TOKELDAALNELGPDGDEEE 80
        |||||
DB      137 TOKELDAAPNELGPDGDEEE 156
        |||||

RESULT 12
US-08-529-055-47
; Sequence 47, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yoether, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; QUERY MATCH          90.2%; Score 358; DB 3; Length 641;
; BEST LOCAL SIMILARITY 90.2%; Score 358; DB 3; Length 641;
; MATCHES 74; CONSERVATIVE 1; MISMATCHES 5; INDELS 0; GAPS 0;

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 51.102 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-19

Perfect score: 397

Sequence: 1 LKKEAGEALDKKADGLPNK.....TKELDAALNELGPDGDEEE 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pap:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pap:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pap:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pap:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pap:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pap:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pap:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pap:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pap:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pap:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pap:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pap:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pap:\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pap:\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 397   | 100.0       | 80     | 15 | US-10-674-755-19  |
| 2          | 397   | 100.0       | 197    | 15 | US-10-299-636-59  |
| 3          | 378   | 95.2        | 104    | 15 | US-10-674-755-20  |
| 4          | 375   | 94.5        | 104    | 15 | US-10-674-755-21  |
| 5          | 369.5 | 92.1        | 102    | 15 | US-10-674-755-18  |
| 6          | 368   | 93.7        | 233    | 15 | US-10-299-636-67  |
| 7          | 365   | 91.9        | 213    | 15 | US-10-299-636-62  |
| 8          | 362   | 91.2        | 744    | 10 | US-09-769-787-184 |
| 9          | 362   | 91.2        | 744    | 17 | US-10-472-928-32  |
| 10         | 358   | 90.2        | 641    | 9  | US-09-765-272-160 |
| 11         | 238   | 59.9        | 108    | 15 | US-10-674-755-23  |
| 12         | 238   | 59.9        | 238    | 15 | US-10-299-636-82  |
| 13         | 235   | 59.2        | 235    | 15 | US-10-299-636-83  |
| 14         | 233   | 58.7        | 233    | 15 | US-10-299-636-85  |
| 15         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 16         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 17         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 18         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 19         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 20         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 21         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 22         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 23         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 24         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 25         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 26         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 27         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 28         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 29         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 30         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 31         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 32         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 33         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 34         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 35         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 36         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 37         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 38         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 39         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 40         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 41         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 42         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 43         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 44         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |
| 45         | 233   | 58.7        | 233    | 15 | US-10-299-636-88  |

|    |       |      |     |    |                      |                    |
|----|-------|------|-----|----|----------------------|--------------------|
| 12 | 238   | 59.9 | 211 | 15 | US-10-299-636-82     | Sequence 82, Appl  |
| 13 | 235   | 59.2 | 108 | 15 | US-10-674-755-24     | Sequence 24, Appl  |
| 14 | 235   | 59.2 | 212 | 15 | US-10-299-636-83     | Sequence 83, Appl  |
| 15 | 233   | 58.7 | 108 | 15 | US-10-674-755-25     | Sequence 25, Appl  |
| 16 | 233   | 58.7 | 232 | 15 | US-10-299-636-85     | Sequence 85, Appl  |
| 17 | 233   | 58.7 | 275 | 16 | US-10-414-532-1      | Sequence 1, Appl   |
| 18 | 233   | 58.7 | 458 | 15 | US-10-299-636-88     | Sequence 88, Appl  |
| 19 | 233   | 58.7 | 653 | 16 | US-10-414-532-26     | Sequence 26, Appl  |
| 20 | 235   | 56.7 | 106 | 15 | US-10-674-755-22     | Sequence 22, Appl  |
| 21 | 186   | 46.9 | 108 | 15 | US-10-674-755-26     | Sequence 26, Appl  |
| 22 | 186   | 46.9 | 185 | 15 | US-10-299-636-84     | Sequence 84, Appl  |
| 23 | 149.5 | 37.7 | 119 | 15 | US-10-674-755-27     | Sequence 27, Appl  |
| 24 | 149.5 | 37.7 | 215 | 15 | US-10-299-636-58     | Sequence 58, Appl  |
| 25 | 149   | 37.5 | 459 | 16 | US-10-702-305A-18    | Sequence 18, Appl  |
| 26 | 118   | 29.7 | 487 | 16 | US-10-414-532-34     | Sequence 34, Appl  |
| 27 | 118   | 29.7 | 487 | 16 | US-10-414-532-21     | Sequence 21, Appl  |
| 28 | 118   | 29.7 | 524 | 16 | US-10-414-532-28     | Sequence 28, Appl  |
| 29 | 109.5 | 27.6 | 204 | 15 | US-10-299-636-66     | Sequence 66, Appl  |
| 30 | 106.5 | 26.8 | 195 | 15 | US-10-299-636-86     | Sequence 86, Appl  |
| 31 | 105.5 | 26.6 | 99  | 15 | US-10-674-755-11     | Sequence 11, Appl  |
| 32 | 105.5 | 26.6 | 198 | 15 | US-10-299-636-76     | Sequence 76, Appl  |
| 33 | 105.5 | 26.6 | 290 | 16 | US-10-414-532-65     | Sequence 65, Appl  |
| 34 | 105.5 | 26.6 | 354 | 15 | US-10-299-636-105    | Sequence 105, Appl |
| 35 | 105.5 | 26.6 | 588 | 15 | US-10-299-636-96     | Sequence 96, Appl  |
| 36 | 105.5 | 26.6 | 619 | 10 | US-09-882-774-1      | Sequence 1, Appl   |
| 37 | 105.5 | 26.6 | 619 | 15 | US-10-282-122A-73702 | Sequence 73702, A  |
| 38 | 105.5 | 26.6 | 619 | 16 | US-10-414-532-72     | Sequence 72, Appl  |
| 39 | 105   | 26.4 | 100 | 15 | US-10-674-755-12     | Sequence 12, Appl  |
| 40 | 104   | 26.2 | 230 | 16 | US-10-414-532-32     | Sequence 32, Appl  |
| 41 | 104   | 26.2 | 230 | 16 | US-10-414-532-19     | Sequence 19, Appl  |
| 42 | 104   | 26.2 | 336 | 15 | US-10-299-636-103    | Sequence 103, Appl |
| 43 | 102   | 25.7 | 99  | 15 | US-10-674-755-13     | Sequence 13, Appl  |
| 44 | 100   | 25.2 | 100 | 15 | US-10-674-755-3      | Sequence 3, Appl   |
| 45 | 99.5  | 25.1 | 99  | 15 | US-10-674-755-15     | Sequence 15, Appl  |

ALIGNMENTS

RESULT 1  
US-10-674-755-19  
; Sequence 19, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; PRIOR FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-19

Query Match 100.0%; Score 397; DB 15; Length 80;  
Best Local Similarity 100.0%; Pred. No. 6.6e-32;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |    |  |    |
|----|----|--|----|
| QY | 1  | LKKEAGEALDKKADGLPNKVSDELEKISNLEILGGADSEDDTAALPNKATKKALEK | 60 |
| DB | 1  | LKKEAGEALDKKADGLPNKVSDELEKISNLEILGGADSEDDTAALPNKATKKALEK | 60 |
| QY | 61 | TKELDAALNELGPDGDEEE                                      | 80 |
| DB | 61 | TKELDAALNELGPDGDEEE                                      | 80 |

RESULT 2

```
US-10-299-636-59
; Sequence 59, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-59

Query Match          100.0%; Score 397; DB 15; Length 197;
Best Local Similarity 100.0%; Pred. No. 2e-31;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
Db 46 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 105

Qy 61 TQKELDAALNELGPDGDEEE 80
Db 106 TQKELDAALNELGPDGDEEE 125

RESULT 3
US-10-674-755-20
; Sequence 20, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-20

Query Match          95.2%; Score 378; DB 15; Length 104;
Best Local Similarity 96.2%; Pred. No. 7e-30;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 84

Qy 61 TQKELDAALNELGPDGDEEE 80
Db 85 TQKELDAALNELGPDGDEEE 104

US-10-299-636-67
; Sequence 67, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
```

```
US-10-674-755-21
; Sequence 21, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-21

Query Match          94.5%; Score 375; DB 15; Length 104;
Best Local Similarity 95.0%; Pred. No. 1.4e-29;
Matches 76; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
Db 25 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 84

Qy 61 TQKELDAALNELGPDGDEEE 80
Db 85 TQKELDAALNELGPDGDEEE 104

RESULT 5
US-10-674-755-18
; Sequence 18, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-18

Query Match          93.1%; Score 369.5; DB 15; Length 102;
Best Local Similarity 96.2%; Pred. No. 4.7e-29;
Matches 77; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
Db 24 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 82

Qy 61 TQKELDAALNELGPDGDEEE 80
Db 83 TQKELDAALNELGPDGDEEE 102

RESULT 6
US-10-299-636-67
; Sequence 67, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
```

```
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; OTHER INFORMATION: Xaa at position 129 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (131)
; OTHER INFORMATION: Xaa at position 131 is unknown
; US-10-299-636-67

Query Match          92.7%; Score 368; DB 15; Length 233;
Best Local Similarity 93.8%; Pred. No. 1.9e-28;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 77 LDKAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 136
QY 61 TKELDAALNELGPDGDEEE 80
DB 137 TKELDAALNELGPDGDEEE 156

RESULT 7
US-10-299-636-62
; Sequence 62, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
```

```
; LOCATION: (2)
; OTHER INFORMATION: Xaa at position 2 is unknown
; US-10-299-636-62

Query Match          91.9%; Score 365; DB 15; Length 213;
Best Local Similarity 92.5%; Pred. No. 3.3e-28;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDKAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 83 LDKAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 142
QY 61 TKELDAALNELGPDGDEEE 80
DB 143 TKELDAALNELGPDGDEEE 162

RESULT 8
US-09-769-787-184
; Sequence 184, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 184
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-769-787-184

Query Match          91.2%; Score 362; DB 10; Length 744;
Best Local Similarity 93.8%; Pred. No. 3.1e-27;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDKAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
DB 370 LDKAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 429
QY 61 TKELDAALNELGPDGDEEE 80
DB 430 TKELDAALNELGPDGDEEE 449

RESULT 9
US-10-472-928-32
; Sequence 32, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 32
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
```





Search completed: June 21, 2005, 11:18:37  
Job time : 51.102 secs

```

RESULT 14
US-10-299-636-83
; Sequence 83, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-83

```

```

Query Match      59.2%; Score 235; DB 15; Length 212;
Best Local Similarity 63.1%; Pred. No. 2.5e-15;
Matches 53; Conservative 11; Mismatches 16; Indels 4; Gaps 2;
QV 1 LDKEAGBAEILDKKAGLGNPKVSDLEKISNLEILLGGADS---ED-DTAALPNKLATKKA 56

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 8 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-19  
Perfect score: 397  
Sequence: 1 LDKEAGEALDKKADGLPNK.....TQKELDAALNELGPDGDEEE 80

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 362   | 91.2        | 744    | 2 F95013 | pneumococcal surfa |
| 2          | 105.5 | 26.6        | 619    | 2 A97887 | surface protein ps |
| 3          | 105.5 | 26.6        | 619    | 2 A41971 | surface protein ps |
| 4          | 89.5  | 22.5        | 284    | 2 S19691 | tropomyosin alpha, |
| 5          | 88.5  | 22.3        | 284    | 2 JC6199 | alpha-tropomyosin  |
| 6          | 87    | 21.9        | 415    | 2 S35760 | fcrA protein precu |
| 7          | 86.5  | 21.8        | 284    | 2 A45488 | body-wall muscle t |
| 8          | 85.5  | 21.5        | 284    | 2 JC6198 | alpha-tropomyosin  |
| 9          | 85    | 21.4        | 405    | 2 A33939 | Fc gamma (19G) rec |
| 10         | 84.5  | 21.3        | 281    | 2 A34787 | tropomyosin 1 alph |
| 11         | 84.5  | 21.3        | 284    | 1 TMRBA  | tropomyosin alpha  |
| 12         | 84.5  | 21.3        | 284    | 2 A39816 | tropomyosin 2, fib |
| 13         | 84.5  | 21.3        | 284    | 2 B27407 | tropomyosin alpha  |
| 14         | 84.5  | 21.3        | 284    | 2 A27674 | tropomyosin 3, fib |
| 15         | 84.5  | 21.3        | 284    | 2 A25825 | tropomyosin alpha  |
| 16         | 84.5  | 21.3        | 284    | 2 JC2551 | tropomyosin alpha  |
| 17         | 84.5  | 21.3        | 284    | 2 A60597 | tropomyosin 2, fib |
| 18         | 84.5  | 21.3        | 284    | 2 B39816 | tropomyosin 3, fib |
| 19         | 84.5  | 21.3        | 2139   | 2 T18296 | myosin heavy chain |
| 20         | 84    | 21.2        | 1509   | 1 A27224 | myosin heavy chain |
| 21         | 83    | 20.9        | 388    | 2 A46173 | Mrp4 protein - Str |
| 22         | 83    | 20.9        | 516    | 2 B84709 | hypothetical prote |
| 23         | 83    | 20.9        | 1837   | 2 T41023 | probable nuclear p |
| 24         | 82.5  | 20.8        | 280    | 2 A22165 | tropomyosin alpha  |
| 25         | 82.5  | 20.8        | 284    | 2 S24972 | tropomyosin alpha, |
| 26         | 82    | 20.7        | 936    | 2 S39083 | myosin heavy chain |
| 27         | 81.5  | 20.5        | 308    | 2 T08796 | tropomyosin - huma |
| 28         | 81.5  | 20.5        | 587    | 2 JC1419 | Fc gamma (19G) rec |
| 29         | 81.5  | 20.5        | 1051   | 2 T18302 | apsB protein - Eme |

ALIGNMENTS

RESULT 1

F95013  
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: F95013  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidt  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf, I.  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: F95013  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-744 <KUR>  
A:Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl4971584; GSPDB:B  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP0117

|                       |       |                       |                                      |                                 |
|-----------------------|-------|-----------------------|--------------------------------------|---------------------------------|
| Query Match           | 91.2% | Score 362;            | DB 2;                                | Length 744;                     |
| Best Local Similarity | 93.8% | Pred. No. 2e-22;      |                                      |                                 |
| Matches               | 75;   | Conservative          | 1;                                   | Mismatches 4; Indels 0; Gaps 0; |
| QY                    | 1     | LDKEAGEALDKKADGLPNKVS | LDKEISNLEILIGGADSEDDTAALPNKLTAKAELEK | 60                              |
| DB                    | 370   | LDKEAGEALDKKADGLPNKVS | LDKEISNLEILIGGADSEDDTAALPNKLTAKAELEK | 429                             |
| QY                    | 61    | TQKELDAALNELGPDGDEEE  | 80                                   |                                 |
| DB                    | 430   | TQKELDAALNELGPDGDEEE  | 449                                  |                                 |

RESULT 2

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>

30 81 20.4 1938 1 A40997 myosin heavy chain  
31 81 20.4 1940 2 A29320 alpha-tropomyosin  
32 80.5 20.3 284 2 I51731 tropomyosin beta -  
33 80.5 20.3 284 2 S23256 probable RNA helic  
34 80.5 20.3 770 2 S56805 paramyosin - nemat  
35 80.5 20.3 848 2 A44972 chromosome segrega  
36 80.5 20.3 1177 2 B75150 kinesin-related pr  
37 80.5 20.3 2954 2 T4156 fcrA 15 protein -  
38 80 20.2 388 2 S52536 embryonic muscle m  
39 80 20.2 1927 2 A59236 myosin heavy chain  
40 80 20.2 1939 2 I48175 paramyosin - nemat  
41 79.5 20.0 879 2 A48575 myosin alpha heavy  
42 79 19.9 234 2 S33732 myosin alpha heavy  
43 79 19.9 465 2 A02986 myosin heavy chain  
44 79 19.9 520 2 S35575 unknown protein, 7  
45 79 19.9 555 2 C96667



A:Residues: 1-415 <POD>  
A:Cross-references: UNIPROT:Q54859; EMBL:X69324; NID:g311759; PIDN:CAA49165.1; PID:g3117  
R:Haanes, E.J.; Heath, D.G.; Cleary, P.P.  
J. Bacteriol. 174, 4967-4976, 1992  
A:Title: Architecture of the vir regulons of group A streptococci parallels opacity factor  
A:Reference number: A42711; MUID:92332431; PMID:1385809  
A:Accession: A42711  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 343-415 <HAA>  
A:Cross-references: GB:M86806; NID:g153630; PIDN:AAA26887.1; PID:g153631  
A:Experimental source: strain CS101, OP+  
A:Note: sequence extracted from NCBI backbone (NCBI:108942, NCBIP:108945)  
C:Superfamily: M5 protein

Query Match 21.9%; Score 87; DB 2; Length 415;  
Best Local Similarity 33.3%; Pred. No. 4.9;  
Matches 28; Conservative 12; Mismatches 28; Indels 16; Gaps 3;

QY 3 KEAGEALDKKADGLPNK---VSDLEKEISNLEILLGG-----ADSEDDTAAL 47  
DB 125 KAAAEAEAKDALNNKQISDLTNEAQLKEAIEGYVQTITQNASRBAIAKQOEAAV 184  
QY 48 PNKLATKKALEKTQKELDAALNE 71  
DB 185 KSQLEAKNAEIDL-KQQDASKTE 207

RESULT 7  
A45488  
body-wall muscle tropomyosin - sea squirt (Ciona intestinalis)  
C:Species: Ciona intestinalis  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A45488  
R:Meedel, T.H.; Hastings, K.E.  
J. Biol. Chem. 268, 6735-6764, 1993  
A:Title: Striated muscle-type tropomyosin in a chordate smooth muscle, ascidian body-wall  
A:Reference number: A45488; MUID:93203280; PMID:8454648  
A:Accession: A45488  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-284 <NEE>  
A:Cross-references: UNIPROT:Q07068; GB:X64105; NID:g297457; PIDN:CAA45469.1; PID:g297458  
A:Note: sequence extracted from NCBI backbone (NCBI:128007, NCBIP:128008)  
C:Superfamily: tropomyosin

Query Match 21.8%; Score 86.5; DB 2; Length 284;  
Best Local Similarity 29.7%; Pred. No. 3.7;  
Matches 30; Conservative 16; Mismatches 26; Indels 29; Gaps 4;

QY 1 LDKE-----AGEALDK-----KADGLPNKVSLEKE-----ISNLE 32  
DB 13 LDKENAIADRAEQARTDKSAEDKATGLBEELQGLKRLKATEDELDTSQEKLRTAENLE 72  
QY 33 -ILGGADSEDDTAALPNKLATKKALEKTQKELDAALNEL 72  
DB 73 NAEKKAADAEQEVASLNRRITVLVEELDRAQERLTISL 113

RESULT 8  
JC6198  
alpha-tropomyosin C-2 - axolotl  
C:Species: Ambystoma mexicanum (axolotl)  
C:Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
C:Accession: JC6198  
R:Luque, E.A.; Spinner, B.J.; Dube, S.; Dube, D.K.; Lemanski, L.F.  
Gene 185, 175-180, 1997  
A:Title: Differential expression of a novel isoform of alpha-tropomyosin in cardiac and  
A:Reference number: JC6198; MUID:97208870; PMID:9055812  
A:Contents: Heart  
A:Accession: JC6198  
A:Molecule type: mRNA  
A:Residues: 1-284 <LUQ>

A:Cross-references: UNIPROT:P87348; GB:U33449; NID:gi871355; PIDN:AAC60091.1; PID:gi87135  
C:Comment: This protein is an actin-binding protein.  
C:Genetics:  
A:Gene: Atmc-2  
C:Superfamily: tropomyosin  
C:Keywords: actin binding

Query Match 21.5%; Score 85.5; DB 2; Length 284;  
Best Local Similarity 29.1%; Pred. No. 4.4;  
Matches 30; Conservative 12; Mismatches 28; Indels 33; Gaps 3;

QY 1 LDKE-----AGEALDKKADGLPNKVSLEKEISNLEILL----- 35  
DB 13 IDKENAMRAEQEADKK--GAEDSKQLESEIVQLEKQIRISEDRDRVLDLHKSEES 70  
QY 36 -----GGADSEDDTAALPNKLATKKALEKTQKELDAALNEL 72  
DB 71 LUTADEKAAKGSDDAASLNRRITQLVEEELDRAQERLATLQKL 113

RESULT 9  
A33939  
Fc gamma (IgG) receptor II precursor - Streptococcus sp. (fragment)  
C:Species: Streptococcus sp.  
C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 26-Aug-1999  
C:Accession: A33939  
R:Heath, D.G.; Cleary, P.P.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4741-4745, 1989  
A:Title: Fc-receptor and M-protein genes of group A streptococci are products of gene duplication  
A:Reference number: A33939; MUID:89282846; PMID:2660147  
A:Accession: A33939  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-405 <HEA>  
A:Cross-references: GB:M22532; NID:g153628; PIDN:AAB95296.1; PID:g552003  
C:Superfamily: M5 protein  
C:Keywords: immunoglobulin receptor

Query Match 21.4%; Score 85; DB 2; Length 405;  
Best Local Similarity 34.2%; Pred. No. 7;  
Matches 25; Conservative 12; Mismatches 28; Indels 8; Gaps 2;

QY 3 KEAGEALDKKADGLPNK---VSDLEKEISNLEILLGG-----ADSEDDTAALPNKLATK 54  
DB 135 KAAAEAEAKDALNNKQISDLTNEAQLKEAIEGYVQTITQNASRBAIAKQOEALAA 194  
QY 55 KAELEKTQKELDA 67  
DB 195 KSQLEAKNAEIEA 207

RESULT 10  
A34787  
tropomyosin 1 alpha, brain - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 09-Jul-2004  
C:Accession: A34787  
R:Lees-Miller, J.P.; Goodwin, L.O.; Helfman, D.M.  
Mol. Cell. Biol. 10, 1729-1742, 1990  
A:Title: Three novel brain tropomyosin isoforms are expressed from the rat alpha-tropomyosin  
A:Reference number: A34787; MUID:90205854; PMID:2320008  
A:Accession: A34787  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-281 <LEE>  
A:Cross-references: UNIPROT:P04692; GB:M34135  
C:Superfamily: tropomyosin  
C:Keywords: alternative splicing; coiled coil

Query Match 21.3%; Score 84.5; DB 2; Length 281;  
Best Local Similarity 28.2%; Pred. No. 5.3;  
Matches 29; Conservative 15; Mismatches 26; Indels 33; Gaps 3;



Search completed: June 21, 2005, 10:12:04  
Job time : 8 secs

**This Page Blank (uspto)**



```

12/
KN SEQUENCE FROM N.A.
RP STRAIN=60;
RC Beall B.W.;
RA
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RR EMBL: AF253404; AAF67352.1; -.
DR InterPro; IPR009053; PfFoldin.
DN
FT NON_TER 1
FT NON_TER 228
FT NON_TER 228
SQ SEQUENCE 228 AA; 24430 MW; E6EAA953EC54EA0F CRC64;

Query Match 91.2%; Score 362; DB 2; Length 228;
Best Local Similarity 93.8%; Pred. No. 8.3e-22;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LDKAGEAEELDKADGLPNKVSDELEKEINLEILLGGADSEDDTAALPNKLTATKKALEK 60
Db 70 LDKAGEAEELDKADGLPNKVSDELEKEINLEILLGGADSEDDTAALPNKLTATKKALEK 129
Qy 61 TOKELDAALNELGPDGDEE 80
Db 130 TOKELDAALNELGPDGDEE 149

RESULT 2
Q9LS82 PRELIMINARY; PRT; 235 AA.
ID Q9LS82
AC Q9LS82;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PsdpA (Fragment).

```



```
RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL: AY082389; AAL92494.1; -.
DR InterPro: IPR000533; Tropomyosin.
DR PRINTS: PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 360
SQ SEQUENCE 360 AA; 39575 MW; 0C09A791547A47EC CRC64;

Query Match 91.2%; Score 362; DB 2; Length 360;
Best Local Similarity 93.8%; Pred. No. 1.3e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
DB 256 LKKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 315
QY 61 TQKELDAALNELGPDGDEEE 80
DB 316 TQKELDAALNELGPDGDEEE 335

RESULT 6
Q9LAX7 PRELIMINARY; PRT; 429 AA.
AC Q9LAX7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AC122;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071816; AAF27714.1; -.
DR HSSP: P04268; IIC2.
DR InterPro: IPR011047; Quin_abc_DH_like.
FT NON_TER 526
SQ SEQUENCE 429 AA; 47048 MW; BC1D74BBAS4DA9D6 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 429;
Best Local Similarity 93.8%; Pred. No. 1.6e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
DB 278 LKKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 337
QY 61 TQKELDAALNELGPDGDEEE 80
DB 338 TQKELDAALNELGPDGDEEE 357

RESULT 7
Q9LAX9 PRELIMINARY; PRT; 526 AA.
AC Q9LAX9;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AC122;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071818; AAF27714.1; -.
DR InterPro: IPR009053; Prefoldin.
DR PRINTS: PR00194; TROPOMYOSIN.
FT NON_TER 429
SQ SEQUENCE 429 AA; 47048 MW; BC1D74BBAS4DA9D6 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 429;
Best Local Similarity 93.8%; Pred. No. 1.6e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
DB 278 LKKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 337
QY 61 TQKELDAALNELGPDGDEEE 80
DB 338 TQKELDAALNELGPDGDEEE 357
```

```
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF3296;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL: AF071816; AAF27712.1; -.
DR HSSP: P04268; IIC2.
DR InterPro: IPR011047; Quin_abc_DH_like.
FT NON_TER 526
SQ SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AE CRC64;

Query Match 91.2%; Score 362; DB 2; Length 526;
Best Local Similarity 93.8%; Pred. No. 2e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
DB 370 LKKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 429
QY 61 TQKELDAALNELGPDGDEEE 80
DB 430 TQKELDAALNELGPDGDEEE 449

RESULT 8
Q8VQ55 PRELIMINARY; PRT; 608 AA.
AC Q8VQ55;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNIH1156;
RA Lee K.J., Bae S.W., Chung K.S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF460993; AAL67804.1; -.
DR HSSP: P06653; IHCX.
DR Pfam: PF01473; CW binding 1; 10.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_2.
FT NON_TER 608
SQ SEQUENCE 608 AA; 67918 MW; 15F71BD62E297526 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 608;
Best Local Similarity 93.8%; Pred. No. 2.3e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKKEAGEAELDKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
DB 246 LKKEAEAEELDKADELQNKVADLEKEISNLEILLGGADSEDDTAALQNKLTAKKAELEK 305
QY 61 TQKELDAALNELGPDGDEEE 80
```

```
Db 306 TQKELDAALNELGPDGDEE 325
RESULT 9
Q97T39 PRELIMINARY; PRT; 744 AA.
AC Q97T39;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A.
GN OrderedLocusNames=SP0117;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang P., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506 (2001).
DR EMBL; AE007328; AAK74303.1; -.
DR PIR; F95013; F95013.
DR HSP; P06653; IHXC.
DR TIGR; SP0117; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN 2.
KW Complete proteome.
SQ SEQUENCE 744 AA; 82764 MW; 20EASE8E7911EFD5 CRC64;

Query Match 91.2%; Score 362; DB 2; Length 744;
Best Local Similarity 93.8%; Pred. No. 2.9e-21;
Matches 75; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 370 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 429
Qy 61 TQKELDAALNELGPDGDEE 80
Db 430 TQKELDAALNELGPDGDEE 449
Qy 61 TQKELDAALNELGPDGDEE 80
Db 430 TQKELDAALNELGPDGDEE 449

RESULT 10
Q9L579 PRELIMINARY; PRT; 231 AA.
AC Q9L579;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255547; AAF68100.1; -.
DR HSP; P04268; IIC2.
FT NON_TER 1
FT NON_TER 241
SQ SEQUENCE 241 AA; 26038 MW; BB87E1A4C25FA669 CRC64;

Query Match 90.4%; Score 359; DB 2; Length 241;
Best Local Similarity 92.5%; Pred. No. 1.5e-21;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 101 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 160
Qy 61 TQKELDAALNELGPDGDEE 80
Db 161 TQKELDAALNELGPDGDEE 180

"pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255547; AAF68100.1; -.
DR HSP; P04268; IIC2.
FT NON_TER 1
FT NON_TER 241
SQ SEQUENCE 241 AA; 24990 MW; A7731F3A46460186 CRC64;

Query Match 90.4%; Score 359; DB 2; Length 231;
Best Local Similarity 92.5%; Pred. No. 1.5e-21;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 100 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 159
Qy 61 TQKELDAALNELGPDGDEE 80
Db 160 TQKELDAALNELGPDGDEE 179

RESULT 11
Q9L580 PRELIMINARY; PRT; 241 AA.
AC Q9L580;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255547; AAF68100.1; -.
DR HSP; P04268; IIC2.
FT NON_TER 1
FT NON_TER 241
SQ SEQUENCE 241 AA; 26038 MW; BB87E1A4C25FA669 CRC64;

Query Match 90.4%; Score 359; DB 2; Length 241;
Best Local Similarity 92.5%; Pred. No. 1.5e-21;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 101 LDKEAGEAELDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 160
Qy 61 TQKELDAALNELGPDGDEE 80
Db 161 TQKELDAALNELGPDGDEE 180
```

```

RESULT 12
Q9L562
ID Q9L562 PRELIMINARY; PRT; 242 AA.
AC Q9L562;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=69;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=69;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255908; AAF70098.1;
FT NON_TER 1
FT NON_TER 242
SQ SEQUENCE 242 AA; 25843 MW; 707BA930797D2C82 CRC64;

Query Match 90.2%; Score 358; DB 2; Length 242;
Best Local Similarity 91.2%; Pred. No. 1.9e-21;
Matches 73; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 94 LKKEAAEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 153
QY 61 TKELDAALNELGPDGDEE 80
Db 154 TKELDAALNELGPDGDEE 173

RESULT 13
Q9L5B7
ID Q9L5B7 PRELIMINARY; PRT; 249 AA.
AC Q9L5B7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RESULT 14
Q9LAX8
ID Q9LAX8 PRELIMINARY; PRT; 502 AA.
AC Q9LAX8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8090;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of pspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae."
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071817; AAF27713.1;
DR HSSP; O15813; 1D7M.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PRO0194; TROPOMYOSIN.
FT NON_TER 502
SQ SEQUENCE 502 AA; 55018 MW; 4E73D477CAE79B40 CRC64;

Query Match 89.4%; Score 355; DB 2; Length 502;
Best Local Similarity 91.2%; Pred. No. 7e-21;
Matches 73; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKKEAGEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 60
Db 371 LKKEAAEALDKKADGLPNKVSDEKEISNLEILLGGADSEDDTAALPNKLTAKKAELEK 430
QY 61 TKELDAALNELGPDGDEE 80
Db 431 TKELDAALNELGPDGDEE 450

RESULT 15
Q9L593
ID Q9L593 PRELIMINARY; PRT; 209 AA.
AC Q9L593;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PepA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```

```

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=115;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=115;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254256; AAF68091.1; -.
FT NON_TER 1
FT NON_TER 209
SQ SEQUENCE 209 AA; 22628 MW; 06FF588F7C3BD5B7 CRC64;

Query Match      88.9%; Score 353; DB 2; Length 209;
Best Local Similarity 91.2%; Pred.No. 4.1e-21;
Matches 73; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LDKEAGEAELDKKADGLPNKVSLEKEISNLEILLGGADSEDDTAALPNKLTATKKALEK 60
Db 43 LDKEAEEAELDKKADGLQNKVADLEKEISNLEILLGGADPEDDTAALQNKLTATKKALEK 102

Qy 61 TQKELDAALNELGPDGDEEE 80
Db 103 TQKELDAALNELGPDGDEEE 122

```

Search completed: June 21, 2005, 10:22:14  
 Job time : 50.551 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:50:39 ; Search time 77.5755 Seconds  
(without alignments)  
518.502 Million cell updates/sec

Title: US-10-674-755-20

Perfect score: 514

Sequence: 1 LAKQTELEKLLDSLDPEGK.....TQKELDAALNRLGPDGDEE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 498   | 96.9        | 416    | 8     | Adk52498 alpha hel  |
| 2          | 498   | 96.9        | 526    | 8     | Adk52497 PepA mole  |
| 3          | 498   | 96.9        | 744    | 6     | Abu00449 S. pneumo  |
| 4          | 498   | 96.9        | 744    | 8     | Adm92054 S. pneumon |
| 5          | 498   | 96.9        | 745    | 3     | Aay81652 Streptoco  |
| 6          | 494   | 96.1        | 641    | 2     | Aaw61217 Streptoco  |
| 7          | 494   | 96.1        | 641    | 5     | Abp54636 S. pneumo  |
| 8          | 494   | 96.1        | 641    | 7     | Adc45241 S. pneumo  |
| 9          | 490   | 95.3        | 213    | 7     | Abw02601 Bg8090c p  |
| 10         | 490   | 95.3        | 8991   | 6     | Abu08487 S. pneumo  |
| 11         | 488   | 94.9        | 213    | 2     | Aaw14567 Streptoco  |
| 12         | 486   | 94.6        | 197    | 7     | Abw02598 Ac122c pn  |
| 13         | 485   | 90.5        | 233    | 7     | Abw02606 Bf1019c p  |
| 14         | 462.5 | 90.0        | 196    | 2     | Aaw14564 Streptoco  |
| 15         | 458   | 89.1        | 233    | 2     | Aaw14572 Streptoco  |
| 16         | 319   | 62.1        | 211    | 7     | Abw02621 Bg11703c   |
| 17         | 319   | 62.1        | 238    | 2     | Aaw14587 Streptoco  |
| 18         | 313   | 60.9        | 232    | 7     | Abw02624 Bf5668c p  |
| 19         | 313   | 60.9        | 275    | 8     | Ado52055 S. pneumo  |
| 20         | 313   | 60.9        | 369    | 8     | Adk52496 alpha hel  |
| 21         | 313   | 60.9        | 458    | 2     | Aaw14592 Streptoco  |
| 22         | 313   | 60.9        | 458    | 7     | Abw02626 Bf5668 pn  |
| 23         | 313   | 60.9        | 653    | 8     | Adk52495 PepA mole  |
| 24         | 313   | 60.9        | 653    | 8     | Ado52080 S. pneumo  |
| 25         | 304   | 59.1        | 212    | 2     | Aaw14588 Streptoco  |

|    |       |      |     |   |          |                    |
|----|-------|------|-----|---|----------|--------------------|
| 26 | 304   | 59.1 | 212 | 7 | ABW02622 | Bg7817c p          |
| 27 | 302.5 | 58.9 | 233 | 2 | AAW14590 | Streptoco          |
| 28 | 277   | 53.9 | 459 | 8 | ADO15316 | S. pneumon         |
| 29 | 256   | 49.8 | 185 | 7 | ABW02623 | Bg7561c p          |
| 30 | 242.5 | 47.2 | 184 | 2 | AAW14589 | Streptoco          |
| 31 | 198   | 38.5 | 487 | 8 | ADR04321 | Streptoco          |
| 32 | 198   | 38.5 | 489 | 8 | ADO52088 | Streptoco          |
| 33 | 198   | 38.5 | 524 | 8 | ADO52082 | E. coli B          |
| 34 | 198   | 38.5 | 627 | 8 | ADO52129 | Ado52129 E. coli B |
| 35 | 193.5 | 37.6 | 119 | 2 | AAW46291 | Pneumococ          |
| 36 | 193.5 | 37.6 | 215 | 7 | AAW14563 | Streptoco          |
| 37 | 193.5 | 37.6 | 215 | 7 | ABW02597 | Atcc6303c          |
| 38 | 185.5 | 36.1 | 290 | 8 | ADO52119 | Ado52119 PYA3637 b |
| 39 | 185.5 | 36.1 | 298 | 8 | ADO52127 | Ado52127 PYA3637 b |
| 40 | 180   | 35.0 | 230 | 8 | ADO52086 | S. pneumo          |
| 41 | 180   | 35.0 | 230 | 8 | ADR04319 | Streptoco          |
| 42 | 129.5 | 25.2 | 550 | 8 | ADK48356 | Streptoco          |
| 43 | 129.5 | 25.2 | 550 | 8 | ADR95223 | Novel S.           |
| 44 | 128.5 | 25.0 | 289 | 2 | AAW62276 | Streptoco          |
| 45 | 128.5 | 25.0 | 289 | 2 | AAW41840 | Streptoco          |

ALIGNMENTS

RESULT 1

ADK52498

ID ADK52498 standard; protein; 416 AA.

XX ADK52498;

XX 20-MAY-2004 (first entry)

XX alpha helical region PspA molecule from the EF3296 strain.

XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;  
KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;  
KW Hodgkin's disease.

XX Streptococcus pneumoniae.

XX WO2004016231-A2.

XX 26-FEB-2004.

XX 17-FEB-2003; 2003WO-US008199.

XX 15-MAR-2002; 2002US-0365351P.

XX (UABR-) UAB RES FOUND.

XX Briles DE;

XX WPI; 2004-192068/18.

XX Treating Streptococcus pneumoniae infection in a subject lacking a functional spleen comprises administering an antibody that recognizes pneumococcal surface protein A (PspA) or its binding portion.

XX Claim 17; SEQ ID NO 4; 41pp; English.

XX The present invention relates to treating Streptococcus pneumoniae infection in a subject lacking a functional spleen comprises administering an antibody that recognizes pneumococcal surface protein A (PspA) or its binding portion. The method is useful for treating or preventing Streptococcus pneumoniae infection in a subject lacking a functional spleen. The disease-associated injury is especially due to hemolytic anemia disease, leukemia or lymphoma, especially sickle cell anemia or Hodgkin's disease. The present sequence represents the alpha helical region PspA molecule from the EF3296 strain of Streptococcus pneumoniae.

XX Sequence 416 AA;

SQ

Query Match 96.9%; Score 498; DB 8; Length 416;  
 Best Local Similarity 98.1%; Pred. No. 1.9e-35;  
 Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKGTQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 60  
 DB 241 LAKKQTELEKLLDLSLDPGKGTQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 300

QY 61 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEEE 104  
 DB 301 GADSEDDTAALQNLKATKKALEKTQKELDAALNELGPDGDEEE 344

RESULT 2  
 ADK52497  
 ID ADK52497 standard; protein; 526 AA.  
 AC ADK52497;  
 DT 20-MAY-2004 (first entry)  
 XX PepA molecule from the EF3296 strain of Streptococcus pneumoniae.  
 DE Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;  
 KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;  
 KW Hodgkin's disease.  
 XX Streptococcus pneumoniae.  
 OS WO20040162311-A2.  
 PN 26-FEB-2004.  
 PD 17-FEB-2003; 2003WO-US008199.  
 PF 15-MAR-2002; 2002US-0365351P.  
 PR (UABR-) UAB RES FOUND.  
 XX Briles DE;  
 PI WPI; 2004-192068/18.  
 DR Treating Streptococcus pneumoniae infection in a subject lacking a  
 PT functional spleen comprises administering an antibody that recognizes  
 PT pneumococcal surface protein A (PspA) or its binding portion.  
 XX Claim 17; SEQ ID NO 3; 41pp; English.  
 PS The present invention relates to treating Streptococcus pneumoniae  
 CC infection in a subject lacking a functional spleen comprises  
 CC administering an antibody that recognizes pneumococcal surface protein A  
 CC (PspA) or its binding portion. The method is useful for treating or  
 CC preventing Streptococcus pneumoniae infection in a subject lacking a  
 CC functional spleen. The disease-associated injury is especially due to  
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell  
 CC anemia or Hodgkin's disease. The present sequence represents PspA  
 CC molecule from the EF3296 strain of Streptococcus pneumoniae.  
 XX Sequence 526 AA;  
 SQ

Query Match 96.9%; Score 498; DB 8; Length 526;  
 Best Local Similarity 98.1%; Pred. No. 2.4e-35;  
 Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKGTQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 60  
 DB 346 LAKKQTELEKLLDLSLDPGKGTQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 405

QY 61 GADSEDDTAALPNKLTAKKAELEKTQKELDAALNELGPDGDEEE 104  
 DB 406 GADSEDDTAALQNLKATKKALEKTQKELDAALNELGPDGDEEE 449

RESULT 3  
 ABU00449  
 ID ABU00449 standard; protein; 744 AA.  
 XX ABU00449;  
 AC 23-OCT-2003 (revised)  
 DT 11-FEB-2003 (first entry)  
 XX S. pneumoniae type 4 strain protein from coding region #16.  
 DE Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
 KW gene therapy; vaccine.  
 XX Streptococcus pneumoniae; type 4 strain.  
 OS WO200277021-A2.  
 PN 03-OCT-2002.  
 PD 27-MAR-2002; 2002WO-IB002163.  
 PF 27-MAR-2001; 2001GB-00007658.  
 PR (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX Masignani V, Tettelin H, Fraser C;  
 WPI; 2003-040579/03.  
 DR N-PSDB; ABX05728.  
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
 PT useful as medicaments for treating or preventing a disease or infection  
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or  
 PT ear infection.  
 XX Claim 1; SEQ ID NO 32; 56pp; English.  
 PS The invention relates to a protein comprising or having at least 50%  
 CC identity to any of the 2469 amino acid sequences, identified in the  
 CC specification (available on a computer readable format), or its fragment,  
 CC expressed from 2469 of 2489 identified DNA coding regions from the  
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as  
 CC ABS56454. Also included are an antibody which binds one of the proteins,  
 CC treating a patient by administering the protein, DNA or antibody (in a  
 CC composition), a kit comprising first and second primers, which are the  
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a  
 CC sequence not defined in the specification, for amplifying a target  
 CC sequence contained within a Streptococcus nucleic acid sequence, where  
 CC the first primer is substantially complementary to the target sequence,  
 CC and the second primer is substantially complementary to the complement of  
 CC the target sequence, and where the parts of the primers having  
 CC substantial complementarity define the termini of the target sequence to  
 CC be amplified, assay comprising contacting a test compound with the  
 CC protein, and determining whether the test compound binds to the protein  
 CC and a Streptococcus pneumoniae bacterium, where one or more genes  
 CC encoding the proteins has been rendered inactive. The proteins, nucleic  
 CC acid molecules, antibody and compositions are useful as medicaments for  
 CC treating or preventing a disease or infection due to streptococcus  
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis  
 CC media or ear infection. They are also useful in developing vaccines,  
 CC diagnostics and antibiotics. The methods are useful for identifying  
 CC immunodominant proteins. The present sequence is one of the 2469 proteins  
 CC expressed by the identified coding regions from the genomic sequence.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 23-Oct-2003 to  
 CC standardise OS field)  
 XX



```
SQ Sequence 744 AA;
Query Match 96.9%; Score 498; DB 6; Length 744;
Best Local Similarity 98.1%; Pred. No. 3.6e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELKKAEEAEELDKKADELQKQVADLEKEISNLEILLG 60
DB 346 LAKKQTELEKLLDSDPEGKTQDELKKAEEAEELDKKADELQKQVADLEKEISNLEILLG 405

QY 61 GADSEDDTAAALPNKLTAKAELEKTQKELDAALNELGPDGDEE 104
DB 406 GADSEDDTAAALQNKLTAKAELEKTQKELDAALNELGPDGDEE 449

RESULT 4
ADM92054
ID ADM92054 standard; protein; 744 AA.
XX
AC ADM92054;
XX
DT 03-JUN-2004 (first entry)
XX
DE S pneumoniae antigenic protein sequence SeqID251.
XX
KW antibacterial; gene therapy; Streptococcus pneumoniae infection;
KW antigenic.
XX
OS Streptococcus pneumoniae.
XX
PN WO2004020609-A2.
XX
PD 11-MAR-2004.
XX
PF 02-SEP-2003; 2003WO-US027401.
XX
PR 30-AUG-2002; 2002US-0407082P.
XX
PA (TUFT ) UNIV TUFTS.
XX
PI Camilli A, Hava DL;
XX
WPI; 2004-239189/22.
DR N-PSDB; ADM91817.
XX
New Streptococcus pneumoniae nucleic acid molecules, useful for
PT diagnosing, treating and preventing active infections of Streptococcus
PT pneumoniae.
XX
PS Claim 27; SEQ ID NO 251; 123pp; English.
XX
This invention relates to novel isolated Streptococcus pneumoniae nucleic
CC acid molecules and the antigenic polypeptides encoded by them. The
CC invention may be useful for the production of compounds with an
CC antibacterial activity or for gene therapy. The nucleic acid molecules,
CC compositions and methods disclosed are useful for treating Streptococcus
CC pneumoniae infection. The present sequence is that of an S pneumoniae
CC protein of the invention.
XX
SQ Sequence 744 AA;
Query Match 96.9%; Score 498; DB 8; Length 744;
Best Local Similarity 98.1%; Pred. No. 3.6e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELKKAEEAEELDKKADELQKQVADLEKEISNLEILLG 60
DB 346 LAKKQTELEKLLDSDPEGKTQDELKKAEEAEELDKKADELQKQVADLEKEISNLEILLG 405

QY 61 GADSEDDTAAALPNKLTAKAELEKTQKELDAALNELGPDGDEE 104
DB 406 GADSEDDTAAALQNKLTAKAELEKTQKELDAALNELGPDGDEE 449

RESULT 5
AA81652
ID AA81652 standard; protein; 745 AA.
XX
AC AA81652;
XX
DT 24-MAY-2000 (first entry)
XX
DE Streptococcus pneumoniae protein sequence ID301.
XX
KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
KW pneumococcal disease.
XX
OS Streptococcus pneumoniae.
XX
PN WO200006737-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-GB002451.
XX
PR 27-JUL-1998; 98GB-00016337.
PR 19-MAR-1999; 99US-0125164P.
XX
PA (MICR-) MICROBIAL TECHNIQS LTD.
XX
PI Gilbert CFG, Hansbro PM;
XX
WPI; 2000-195300/17.
XX
New Streptococcal protein, useful as a vaccine, for diagnosis of
PT pneumococcal diseases and for screening agents capable of antagonizing or
PT inhibiting expression of the protein.
XX
PS Claim 2; Page 95; 108pp; English.
XX
AA81679 represent specifically claimed protein sequences
CC isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC The sequences have antibacterial and antiinflammatory properties. The
CC protein sequences, and fragments of them, are useful as immunogens and/or
CC antigens. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the
CC detection and diagnosis of S. pneumoniae. The protein sequences are also
CC useful for screening an agent capable of antagonising, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC and meningitis. AA05591 to AA05614 represent primers used in the
CC exemplification of the present invention
XX
SQ Sequence 745 AA;
Query Match 96.9%; Score 498; DB 3; Length 745;
Best Local Similarity 98.1%; Pred. No. 3.7e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELKKAEEAEELDKKADELQKQVADLEKEISNLEILLG 60
DB 346 LAKKQTELEKLLDSDPEGKTQDELKKAEEAEELDKKADELQKQVADLEKEISNLEILLG 405

QY 61 GADSEDDTAAALPNKLTAKAELEKTQKELDAALNELGPDGDEE 104
DB 406 GADSEDDTAAALQNKLTAKAELEKTQKELDAALNELGPDGDEE 449

RESULT 6
AA61217
ID AA61217 standard; protein; 641 AA.
XX
AC AA61217;
XX
```

```

DT 02-OCT-1998 (first entry)
DE Streptococcus pneumoniae SP0092 protein.
DE XX Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
DE XX Streptococcus pneumoniae; otitis media; meningitis.
DE XX Streptococcus pneumoniae.
DE XX Key Location/Qualifiers
FT Misc-difference 306 /label= unknown
FT FT /note= "encoded by NCT"
XX WO9818930-A2.
XX 07-MAY-1998.
XX 30-OCT-1997; 97WO-US019422.
XX 31-OCT-1996; 96US-0029960P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunsch CA, Choi GH, Johnson LS, Hromockyj A;
XX PI WPI; 1998-272224/24.
XX DR N-PSDB; AAV27403.
XX Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae
XX - or their epitope-containing fragments, useful in protective or
XX therapeutic vaccines, and for diagnosis.
XX Claim 11; Page 82; 118pp; English.
XX The present sequence represents a protein from Streptococcus pneumoniae.
XX The nucleic acid sequence encoding the Streptococcus pneumoniae protein
XX can be useful in vaccines for inducing protective antibodies against
XX Streptococcus pneumoniae, for treatment or prevention of infection e.g.
XX pneumonia, otitis media or meningitis. Probes based on the nucleic acid
XX are used to detect Streptococcus infection (by usual hybridisation or
XX amplification methods), also for isolating Streptococcus genes or their
XX allelic variants. The protein can be used similarly to detect specific
XX antibodies in standard immunoassays, especially for diagnosing or
XX monitoring infections. Antibodies which bind the protein are used to
XX detect corresponding antigens. To purify the protein and for passive
XX immunisation (optionally coupled to a toxin). Vaccines are administered,
XX e.g. by injection, orally or through the skin, typically at 0.01-1000
XX (especially 10-300) mu g/ml per dose
XX Sequence 641 AA;
XX Query Match 96.1%; Score 494; DB 2; Length 641;
XX Best Local Similarity 97.1%; Pred. No. 6.9e-35;
XX Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX 1 LAKKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADELPKNKADLEKEISNLEILG 60
XX 243 LAKKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADELPKNKADLEKEISNLEILG 302
XX 61 GADSEDDTAALPNKATKKAELKTKQKELDAALNELGPDGDEEE 104
XX 303 GADXEDDTAALQNKATKKAELKTKQKELDAALNELGPDGDEEE 346
XX RESULT 7
XX ABP54636
XX ID ABP54636 standard; protein; 641 AA.
XX AC ABP54636;
XX XX ABP54636;
XX 04-SEP-2002 (first entry)
XX XX

```

---

```

DE S. pneumoniae SP092 protein sequence SEQ ID NO:160.
DE XX Streptococcus pneumoniae; epitope; vaccine; antigenic protein;
DE XX antibacterial; Streptococcal infection; detection.
DE XX Streptococcus pneumoniae.
DE XX US2002061545-A1.
DE XX 23-MAY-2002.
DE XX 22-JAN-2001; 2001US-00765272.
DE XX 30-OCT-1997; 97US-00961083.
DE XX (CHOI/) CHOI G H.
DE XX (KUNS/) KUNSCH C A.
DE XX (BARA/) BARASH S C.
DE XX (DILL/) DILLON P J.
DE XX (DOUG/) DOUGHERTY B.
DE XX (FANN/) FANNON M R.
DE XX (ROSE/) ROSEN C A.
DE XX Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
DE XX Rosen CA;
DE XX WPI; 2002-479261/51.
DE XX N-PSDB; ABQ84871.
DE XX New Streptococcus pneumoniae antigens, useful for detecting Streptococcus
DE XX PT and for preventing or attenuating disease caused by Streptococcus
DE XX PT infection.
DE XX Claim 11; Page 43; 70pp; English.
DE XX ABQ84792 to ABQ84904 represents nucleic acids which encode the
DE XX Streptococcus pneumoniae antigens given in ABP54557 to ABP54669. The S.
DE XX Streptococcus pneumoniae antigens have antibacterial activity and can be used in
DE XX vaccines. The S. pneumoniae antigens can also be used to prevent or
DE XX attenuate a Streptococcal infection in an animal. The polynucleotides
DE XX encoding the S. pneumoniae antigens can be used to detect Streptococcus
DE XX nucleic acids. ABQ84905 to ABQ85130 represent primers used in the cloning
DE XX of S. pneumoniae ORFs (open reading frames) which are used in an example
DE XX from the present invention
DE XX Sequence 641 AA;
DE XX Query Match 96.1%; Score 494; DB 5; Length 641;
DE XX Best Local Similarity 97.1%; Pred. No. 6.9e-35;
DE XX Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DE XX 1 LAKKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADELPKNKADLEKEISNLEILG 60
DE XX 243 LAKKQTELEKLLDSDPEGKTQDELKAEAEAEADKKADELPKNKADLEKEISNLEILG 302
DE XX 61 GADSEDDTAALPNKATKKAELKTKQKELDAALNELGPDGDEEE 104
DE XX 303 GADXEDDTAALQNKATKKAELKTKQKELDAALNELGPDGDEEE 346
DE XX RESULT 8
DE XX ADC45241
DE XX ID ADC45241 standard; protein; 641 AA.
DE XX AC ADC45241;
DE XX XX ADC45241;
DE XX 18-DEC-2003 (first entry)
DE XX S. pneumoniae antigenic protein SP092.
DE XX Antigen; bacterial infection; vaccine; pneumonia; antibacterial.
DE XX Streptococcus pneumoniae.

```



```

PD 31-DEC-2002.
XX
XX PF 16-SEP-1996; 96US-00714741.
XX
XX PR 15-SEP-1995; 95US-00529055.
XX
XX PA (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 2003-361534/34.
XX
XX Isolated PspC amino acid sequence used as polymerase chain reaction or
XX hybridization probe, comprises pneumococcal surface protein having alpha-
XX helical, proline rich and repeat regions.
XX
XX Disclosure; Col 145-189; 186pp; English.
XX
XX The present invention relates to the isolation of Streptococcus
XX pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide
XX sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-
XX like protein having alpha-helical, proline rich and repeat regions. The
XX PspC and PspA proteins may be used in a vaccine to protect against
XX pneumococcal infections. The polynucleotide sequences encoding PspC and
XX PspA may be used for the expression of the proteins, and as PCR primers
XX or hybridisation probes. The present sequence represents S. pneumoniae
XX PspA protein
XX
XX SQ Sequence 8991 AA;

Query Match 95.3%; Score 490; DB 6; Length 8991;
Best Local Similarity 95.2%; Pred. No. 3.2e-33;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLLDLSLDPGKTKQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 60
DB 4625 LAKQTELEKLLDNLDPGKTKQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 4684

QY 61 GADSEDDTAALPNKLTAKAELEKTKOKELDAALNELGPDGDEE 104
DB 4685 GADPEDDTAALPNKLTAKAEFEKTKPELDAALNELGPDGDEE 4728

RESULT 11
AAW14567
ID AAW14567 standard; protein; 213 AA.
XX
XX AC AAW14567;
XX
XX DT 17-OCT-2003 (revised)
XX DT 28-OCT-1997 (first entry)
XX
XX DE Streptococcus pneumoniae PspA central region.
XX
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
XX bacteraemia; pneumonia.
XX
XX OS Streptococcus pneumoniae; strain Bg8090.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 2
XX FT /note= "unidentified amino acid"
XX
XX PN W09709994-A1.
XX
XX PD 20-MAR-1997.
XX
XX PF 16-SEP-1996; 96WO-US014819.
XX
XX PR 15-SEP-1995; 95US-00529055.
XX
XX PA (UABR-) UAB RES FOUND.

XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding

```

```

XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
XX alpha-helix region and some of the proline-rich region, of pneumococcal
XX surface protein A (PspA) of Streptococcus pneumoniae strain Bg8090.
XX Comparison of the N-terminal and central regions (AAW14533-57 and
XX AAW14562-91) of PspA polypeptides from different pneumococcal strains can
XX be used to divide the strains into several families based on sequence
XX homologies. PspA polypeptides, or fragments of them, can be used in
XX vaccines to protect animals against S. pneumoniae infection and hence for
XX the prevention of diseases such as otitis media, meningitis, bacteraemia
XX and pneumonia. The sequence of the 3' half of the PspA alpha-helical
XX region and the immediate 5' tip of the coding sequence are likely to be
XX the critical sequences for predicting PspA cross-reactions and vaccine
XX composition. (Updated on 17-Oct-2003 to standardise OS field)
XX
XX SQ Sequence 213 AA;

Query Match 94.9%; Score 488; DB 2; Length 213;
Best Local Similarity 95.2%; Pred. No. 6.5e-35;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLLDLSLDPGKTKQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 60
DB 59 LAKQTELEKLLDNLDPGKTKQDELKAEAEALDKKADLPKNKADLEKEISNLEILLG 118

QY 61 GADSEDDTAALPNKLTAKAELEKTKOKELDAALNELGPDGDEE 104
DB 119 GADPEDDTAALPNKLTAKAEFEKTKPELDAALNELGPDGDEE 162

RESULT 12
ABW02598
ID ABW02598 standard; protein; 197 AA.
XX
XX AC ABW02598;
XX
XX DT 12-FEB-2004 (first entry)
XX
XX DE Ac122c pneumococcal surface protein A (PspA) central region.
XX
XX KW Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
XX KW immunological; gene therapy; immunostimulant.
XX
XX OS Unidentified.
XX
XX PN US6592876-B1.
XX
XX PD 15-JUL-2003.
XX
XX PF 15-SEP-1995; 95US-00529055.
XX
XX PR 20-APR-1993; 93US-00048896.
XX PR 06-JUN-1995; 95US-00465746.
XX
XX PA (UABR-) UAB RES FOUND.
XX
XX Briles DE, Mcdaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain,
XX comprises at least two different full length isolated gene encoding

```

```

PT pneumococcal surface protein A.
PS Example 6; SEQ ID NO 44; 121pp; English.
XX
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antibodies, an
CC immunological or vaccine compositions, for eliciting antigenic, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Acl22c pneumococcal
CC surface protein A (PspA) central region. This sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 197 AA;

Query Match          94.6%; Score 486; DB 7; Length 197;
Best Local Similarity 95.2%; Pred. No. 8.9e-35;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDSDPEGKTQDELDKAEAEALDKADELPNKVADLEKEISNLEILG 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 22 LAQKQTELGKLDSDPEGKTQDELDKAEAEALDKADELPNKVADLEKEISNLEILG 81
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 82 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 125
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
ID ABW02606 standard; protein; 233 AA.
XX
XX ABW02606;
AC
XX
DT 12-FEB-2004 (first entry)
XX
DE Bfl019c pneumococcal surface protein A (PspA) central region.
XX
XX Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;
KW immunological; gene therapy; immunostimulant.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH Misc-difference 1..233
FT /note= "Xaa = Unknown amino acid"
FT
XX
XX US6592876-B1.
XX
XX 15-JUL-2003.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX 20-APR-1993; 93US-00048896.
XX
XX 06-JUN-1995; 95US-00465746.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;
XX
XX WPI; 2003-862841/80.
XX
XX Immunological composition for obtaining expression products used for
XX detecting the presence of Streptococcus pneumoniae or its strain.
XX comprises at least two different full length isolated gene encoding
XX pneumococcal surface protein A.

```

```

PS Example 6; SEQ ID NO 52; 121pp; English.
XX
XX
CC The present invention relates to an immunological composition comprising
CC at least 2 different full length isolated genes encoding pneumococcal
CC surface protein A (PspA) from different groups based on restriction
CC fragment polymorphism analysis. The invention is useful for obtaining
CC expression products by recombinant techniques to detect, determine,
CC isolate or diagnose the presence of Streptococcus pneumoniae or its
CC strain. The expression product is useful for preparing antibodies, an
CC immunological or vaccine compositions, for eliciting antigenic, an
CC immunological response (other than or additional to antibodies) or a
CC protective response (including antibody or other immunological response
CC by administering compositions to a host). The invention is also useful as
CC vaccines and in gene therapy. The present sequence is Efl019c
CC pneumococcal surface protein A (PspA) central region. This sequence is
XX used in the exemplification of the invention
XX
SQ Sequence 233 AA;

Query Match          90.5%; Score 465; DB 7; Length 233;
Best Local Similarity 92.3%; Pred. No. 7.5e-33;
Matches 96; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDSDPEGKTQDELDKAEAEALDKADELPNKVADLEKEISNLEILG 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 53 LAQKPTGLEKLDSDPEGKTQDELDKAEAEALDKADELPNKVADLEKEISNLEILG 112
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 104
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 113 GADSEDDTAALPNKLTAKAELEKTKQELDAALNELGPDGDEE 156
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
AAW14564
ID AAW14564 standard; protein; 196 AA.
XX
XX AAW14564;
AC
XX
DT 17-OCT-2003 (revised)
DT 28-OCT-1997 (first entry)
XX
XX Streptococcus pneumoniae PspA central region.
XX
XX PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;
KW bacteraemia; pneumonia.
XX
XX Streptococcus pneumoniae; strain Acl22.
XX
XX WO9709994-A1.
XX
XX 20-MAR-1997.
XX
XX 16-SEP-1996; 96WO-US014819.
XX
XX 15-SEP-1995; 95US-00529055.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;
XX Hollingshead S, Tart R, Brooks-Walter A;
XX
XX WPI; 1997-202002/18.
XX
XX Streptococcus pneumoniae surface protein PspC and truncated PspA - used
XX in vaccines for protecting animals against S.pneumoniae infection.
XX
XX Example 6; Fig 13; 296pp; English.
XX
XX This sequence shows the central portion, including the C-terminus of the
XX alpha-helix region and some of the proline-rich region, of pneumococcal
XX surface protein A (PspA) of Streptococcus pneumoniae strain Acl22.
XX Comparison of the N-terminal and central regions (AAW14533-57 and
XX AAW14562-91) of PspA polypeptides from different pneumococcal strains can

```



| Result No. | Score | Query |      | Length | DB                | ID | Description       |
|------------|-------|-------|------|--------|-------------------|----|-------------------|
|            |       | Match | %    |        |                   |    |                   |
| 1          | 514   | 100.0 | 104  | 2      | US-08-710-749-19  |    | Sequence 19, Appl |
| 2          | 514   | 100.0 | 104  | 4      | US-09-147-875A-20 |    | Sequence 20, Appl |
| 3          | 500   | 97.3  | 104  | 4      | US-09-147-875A-21 |    | Sequence 21, Appl |
| 4          | 494   | 96.1  | 104  | 2      | US-08-710-749-20  |    | Sequence 20, Appl |
| 5          | 494   | 96.1  | 641  | 3      | US-08-961-083-160 |    | Sequence 160, App |
| 6          | 494   | 96.1  | 641  | 4      | US-09-536-784-160 |    | Sequence 160, App |
| 7          | 490   | 95.3  | 213  | 4      | US-08-529-055-47  |    | Sequence 47, Appl |
| 8          | 490   | 95.3  | 8991 | 4      | US-08-714-741-32  |    | Sequence 32, Appl |
| 9          | 486   | 94.6  | 197  | 4      | US-08-529-055-44  |    | Sequence 44, Appl |
| 10         | 484   | 94.2  | 102  | 2      | US-08-710-749-21  |    | Sequence 21, Appl |
| 11         | 484   | 94.2  | 102  | 4      | US-09-147-875A-18 |    | Sequence 18, Appl |
| 12         | 465   | 90.5  | 233  | 4      | US-08-529-055-52  |    | Sequence 52, Appl |
| 13         | 378   | 73.5  | 80   | 2      | US-08-710-749-18  |    | Sequence 18, Appl |
| 14         | 378   | 73.5  | 80   | 4      | US-09-147-875A-19 |    | Sequence 19, Appl |
| 15         | 319   | 62.1  | 108  | 2      | US-08-710-749-26  |    | Sequence 26, Appl |
| 16         | 319   | 62.1  | 108  | 4      | US-09-147-875A-23 |    | Sequence 23, Appl |
| 17         | 319   | 62.1  | 211  | 4      | US-08-529-055-67  |    | Sequence 67, Appl |
| 18         | 313   | 60.9  | 108  | 2      | US-08-710-749-24  |    | Sequence 24, Appl |
| 19         | 313   | 60.9  | 108  | 4      | US-09-147-875A-25 |    | Sequence 25, Appl |
| 20         | 313   | 60.9  | 232  | 4      | US-08-529-055-70  |    | Sequence 70, Appl |
| 21         | 313   | 60.9  | 458  | 4      | US-08-529-055-73  |    | Sequence 73, Appl |
| 22         | 311   | 60.5  | 108  | 4      | US-09-147-875A-24 |    | Sequence 24, Appl |
| 23         | 309   | 60.1  | 106  | 4      | US-09-147-875A-22 |    | Sequence 22, Appl |
| 24         | 304   | 59.1  | 212  | 4      | US-08-529-055-68  |    | Sequence 68, Appl |
| 25         | 301   | 58.6  | 108  | 2      | US-08-710-749-22  |    | Sequence 22, Appl |
| 26         | 301   | 58.6  | 108  | 2      | US-08-710-749-23  |    | Sequence 23, Appl |
| 27         | 263   | 51.2  | 108  | 2      | US-08-710-749-25  |    | Sequence 25, Appl |





```
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brookes, A. Anders
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PB340P2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 160:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 641 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-961-083-160

Query Match 96.1%; Score 494; DB 3; Length 641;
Best Local Similarity 97.1%; Pred. No. 1.7e-39;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELDKAEAEELDKKADELPKNKVLADLEKEISNLEILLG 60
DB 243 LAKKQTELEKLLDSDPEGKTQDELDKAEAEELDKKADELPKNKVLADLEKEISNLEILLG 302

QY 61 GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDEE 104
DB 303 GADXEDDTAALQNKLTAKAELEKTOKELDAALNELGPDGDEE 346

RESULT 6
US-09-536-784-160
; Sequence 160, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-536-784-160

Query Match 96.1%; Score 494; DB 4; Length 641;
Best Local Similarity 97.1%; Pred. No. 1.7e-39;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELDKAEAEELDKKADELPKNKVLADLEKEISNLEILLG 60
DB 243 LAKKQTELEKLLDSDPEGKTQDELDKAEAEELDKKADELPKNKVLADLEKEISNLEILLG 302

QY 61 GADSEDDTAALPNKLTAKAELEKTOKELDAALNELGPDGDEE 104
DB 303 GADXEDDTAALQNKLTAKAELEKTOKELDAALNELGPDGDEE 346

RESULT 7
US-08-529-055-47
; Sequence 47, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-529-055-47

Query Match 95.3%; Score 490; DB 4; Length 213;
Best Local Similarity 95.2%; Pred. No. 1.1e-39;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELDKAEAEELDKKADELPKNKVLADLEKEISNLEILLG 60
DB 59 LAKKQTELEKLLDSDPEGKTQDELDKAEAEELDKKADELPKNKVLADLEKEISNLEILLG 118
```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Prommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-21

```

|    | Query Match           | 94.2%   | Score 484;         | DB 2;     | Length 102; |
|----|-----------------------|---|--------------------|-----------|-------------|
|    | Best Local Similarity | 98.1%   | Pred. No. 1.6e-39; |           |             |
|    | Matches 102;          | Conservative 0;   | Mismatches 0;      | Indels 2; | Gaps 2;     |
| Qy | 1                     | LAKQTELEKLLDSEGTQDELDEKAEAEALDKKADELPNKVADLEKEISNLEILLG     | 60                 |           |             |
| Db | 1                     | LAKQTELEKLLD-LDPEGTQDELDEKAE-EAELDKKADELPNKVADLEKEISNLEILLG | 58                 |           |             |
| Qy | 61                    | GAUSEDTPALPNKLTAKKAELEKTQKELDAALNELPGDGEDEE                 | 104                |           |             |
| Db | 59                    | GAUSEDTPALPNKLTAKKAELEKTQKELDAALNELPGDGEDEE                 | 102                |           |             |

```

RESULT 11
US-09-147-875A-18
; Sequence 18, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-18

```

|                           |        |   |           |             |
|---------------------------|--------|---|-----------|-------------|
| Query Match               | 94.2%; | Score 484;  | DB 4;     | Length 102; |
| Best Local Similarity     | 98.1%; | Pred. No. 1.6e-39;  |           |             |
| Matches 102; Conservative | 0;     | Mismatches 0;   | Indels 2; | Gaps 2;     |
| Qy                        | 1      | LAKQTLEKLLDLSLDPEGKTQDELDEKEAREAEELDKKADELPNKVADLEKEISNLEILLG | 60        |             |
| Dd                        | 1      | LAKQTLEKLDD-LDPEGKTQDELDEKA-EAEILDKKADELPNKVADLEKEISNLEILLG   | 58        |             |
| Qy                        | 61     | GAUSEDATAALPNKIATKKAAELETKTELDAALNELPGDGDEE                   | 104       |             |
| Dd                        | 59     | GAUSEDATAALPNKIATKKAAELETKTELDAALNELPGDGDEE                   | 102       |             |

RESULT 12  
US-08-529-055-52  
; Sequence 52, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:

```

1  APPLICANT: Briles, David E.
2  APPLICANT: McDaniel, Larry S.
3  APPLICANT: Swiatlo, Edwin
4  APPLICANT: Yother, Janet
5  APPLICANT: Brooks-Walter, Alexis
6  TITLE OF INVENTION: Pneumococcal Genes, Portions
7  TITLE OF INVENTION: Thereof, Expression Products
8  TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
9  TITLE OF INVENTION: Portions and Products
10 NUMBER OF SEQUENCES: 73
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Curtis, Morris & Safford, P.C.
13 STREET: 530 Fifth Avenue
14 CITY: New York
15 STATE: NY
16 COUNTRY: USA
17 ZIP: 10036
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.30
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/529,055
25 FILING DATE: 15-SEP-1995
26 CLASSIFICATION: 435
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Frommer, William S.
29 REGISTRATION NUMBER: 25,506
30 REFERENCE/DOCKET NUMBER: 454312-2400
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (212) 840-3333
33 TELEFAX: (212) 840-0712
34 INFORMATION FOR SEQ ID NO: 52:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 233 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: peptide
41
42 US-08-529-055-52
43
44 Query Match 90.5%; Score 465; DB 4; Length 233;
45 Best Local Similarity 92.3%; Pred. No. 3e-37;
46 Matches 96; Conservative 1; Mismatches 7; Indels 0
47
48 QY 1 LAKKQTELEKLLDSIDPSGKTKQDELDKAEAEAEELDKKADLPKNKVADLEKEI
49 Db 53 LAQPTGLGKULDSIDPSGKTKQDELDKAEAEAEELDKKADLPKNKVADLEKEI
50
51 QY 61 GADSEDDTAALPNKIATPKAAELEKTKQKELDAALNELGPDGDEE 104
52 Db 113 GADSEDDTAALPNKIATPKAAELEKTKQKELDAALNELGPDGDEE 156
53
54 RESULT 13
55 US-08-710-749-18
56 Sequence 18, Application US/08710749
57 Patent No. 5955089
58 GENERAL INFORMATION:
59 APPLICANT: Briles, David E.
60 APPLICANT: Hollingshead, Susan
61 APPLICANT: Becker, Robert
62 TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
63 TITLE OF INVENTION: PROTEINS
64 NUMBER OF SEQUENCES: 28
65 CORRESPONDENCE ADDRESS:
66 ADDRESSEE: Curtis, Morris & Safford
67 STREET: 530 Fifth Avenue
68 CITY: New York
69 STATE: New York
70 COUNTRY: USA
71 ZIP: 10036

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 10:12:15 ; Search time 66.4327 Seconds  
(without alignments)  
601.118 Million cell updates/sec

Title: US-10-674-755-20

Perfect score: 514

Sequence: 1 LAKKQTELEKLLDLDPEK.....TKELDAALNELGPDGDEE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0

Maximum Match 100%

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA.\*

|     |  |
|-----|--|
| 1:  | /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*  |
| 2:  | /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*   |
| 3:  | /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*  |
| 4:  | /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*  |
| 5:  | /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*  |
| 6:  | /cgn2_6/ptodata/1/pubpaa/PCTU5_PUBCOMB.pep.* |
| 7:  | /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*  |
| 8:  | /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*  |
| 9:  | /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.* |
| 10: | /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.* |
| 11: | /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.* |
| 12: | /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*  |
| 13: | /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.* |
| 14: | /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.* |
| 15: | /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.* |
| 16: | /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.* |
| 17: | /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.* |
| 18: | /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*  |
| 19: | /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.* |
| 20: | /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*  |
| 21: | /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*  |
| 22: | /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 514   | 100.0       | 104    | 15    | US-10-674-755-20  |
| 2          | 500   | 97.3        | 104    | 15    | US-10-674-755-21  |
| 3          | 498   | 96.9        | 744    | 10    | US-09-769-787-184 |
| 4          | 498   | 96.9        | 744    | 17    | US-10-472-928-32  |
| 5          | 494   | 96.1        | 641    | 9     | US-09-765-272-160 |
| 6          | 490   | 95.3        | 213    | 15    | US-10-299-636-62  |
| 7          | 486   | 94.6        | 197    | 15    | US-10-299-636-59  |
| 8          | 484   | 94.2        | 102    | 15    | US-10-674-755-18  |
| 9          | 465   | 90.5        | 233    | 15    | US-10-299-636-67  |
| 10         | 378   | 73.5        | 80     | 15    | US-10-674-755-19  |
| 11         | 319   | 62.1        | 108    | 15    | US-10-674-755-23  |

|    |       |      |     |    |                      |                    |
|----|-------|------|-----|----|----------------------|--------------------|
| 12 | 319   | 62.1 | 211 | 15 | US-10-299-636-82     | Sequence 82, Appl  |
| 13 | 313   | 60.9 | 108 | 15 | US-10-674-755-25     | Sequence 25, Appl  |
| 14 | 313   | 60.9 | 232 | 15 | US-10-299-636-85     | Sequence 85, Appl  |
| 15 | 313   | 60.9 | 275 | 16 | US-10-414-532-1      | Sequence 1, Appl   |
| 16 | 313   | 60.9 | 458 | 15 | US-10-299-636-88     | Sequence 88, Appl  |
| 17 | 313   | 60.9 | 653 | 16 | US-10-414-532-26     | Sequence 26, Appl  |
| 18 | 311   | 60.5 | 108 | 15 | US-10-674-755-24     | Sequence 24, Appl  |
| 19 | 309   | 60.1 | 106 | 15 | US-10-674-755-22     | Sequence 22, Appl  |
| 20 | 304   | 59.1 | 212 | 15 | US-10-299-636-83     | Sequence 83, Appl  |
| 21 | 277   | 53.9 | 459 | 16 | US-10-702-305A-18    | Sequence 18, Appl  |
| 22 | 263   | 51.2 | 108 | 15 | US-10-674-755-26     | Sequence 26, Appl  |
| 23 | 256   | 49.8 | 185 | 15 | US-10-299-636-84     | Sequence 84, Appl  |
| 24 | 198   | 38.5 | 487 | 16 | US-10-414-532-34     | Sequence 34, Appl  |
| 25 | 198   | 38.5 | 487 | 16 | US-10-414-532-21     | Sequence 21, Appl  |
| 26 | 198   | 38.5 | 524 | 16 | US-10-414-532-28     | Sequence 28, Appl  |
| 27 | 193.5 | 37.6 | 119 | 15 | US-10-674-755-27     | Sequence 27, Appl  |
| 28 | 193.5 | 37.6 | 215 | 15 | US-10-299-636-58     | Sequence 58, Appl  |
| 29 | 185.5 | 36.1 | 290 | 16 | US-10-414-532-65     | Sequence 65, Appl  |
| 30 | 180   | 35.0 | 230 | 16 | US-10-414-532-32     | Sequence 32, Appl  |
| 31 | 180   | 35.0 | 230 | 16 | US-10-414-533-19     | Sequence 19, Appl  |
| 32 | 128   | 24.9 | 354 | 15 | US-10-299-636-105    | Sequence 105, Appl |
| 33 | 128   | 24.9 | 588 | 15 | US-10-299-636-96     | Sequence 96, Appl  |
| 34 | 128   | 24.9 | 619 | 10 | US-09-882-774-1      | Sequence 1, Appl   |
| 35 | 128   | 24.9 | 619 | 15 | US-10-282-122A-73702 | Sequence 73702, A  |
| 36 | 128   | 24.9 | 619 | 16 | US-10-414-532-72     | Sequence 72, Appl  |
| 37 | 126   | 24.5 | 141 | 14 | US-10-254-995-2      | Sequence 2, Appl   |
| 38 | 126   | 24.5 | 204 | 15 | US-10-299-636-66     | Sequence 66, Appl  |
| 39 | 126   | 24.5 | 589 | 9  | US-09-748-875-14     | Sequence 14, Appl  |
| 40 | 126   | 24.5 | 589 | 10 | US-09-298-523B-14    | Sequence 14, Appl  |
| 41 | 126   | 24.5 | 589 | 15 | US-10-299-636-97     | Sequence 97, Appl  |
| 42 | 126   | 24.5 | 643 | 15 | US-10-299-636-95     | Sequence 95, Appl  |
| 43 | 126   | 24.5 | 670 | 9  | US-09-748-875-63     | Sequence 63, Appl  |
| 44 | 126   | 24.5 | 670 | 10 | US-09-298-523B-63    | Sequence 63, Appl  |
| 45 | 126   | 24.5 | 690 | 9  | US-09-748-875-61     | Sequence 61, Appl  |

#### ALIGNMENTS

RESULT 1  
US-10-674-755-20  
; Sequence 20, Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; PRIOR FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; US-10-674-755-20

Query Match 100.0%; Score 514; DB 15; Length 104;  
Best Local Similarity 100.0%; Pred. No. 8.8e-36;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |    |  |     |
|----|----|--|-----|
| QY | 1  | LAKKQTELEKLLDLDPEKTDQELDKAEAEELDKKADLPKNKADLEKEISNLEILLG | 60  |
| Pb | 1  | LAKKQTELEKLLDLDPEKTDQELDKAEAEELDKKADLPKNKADLEKEISNLEILLG | 60  |
| QY | 61 | GADSEDDTAALPNKATKKALEKTKQELDAALNELGPDGDEEE               | 104 |
| Pb | 61 | GADSEDDTAALPNKATKKALEKTKQELDAALNELGPDGDEEE               | 104 |

RESULT 2

```
US-10-674-755-21
; Sequence 21, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-21

Query Match          97.3%; Score 500; DB 15; Length 104;
Best Local Similarity 97.1%; Pred. No. 1.3e-34;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
    |||||
Db 1 LAKQTELEKLLDNLDPGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
    |||||

Qy 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104
    |||||
Db 61 GADPEDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104
    |||||

RESULT 3
US-09-769-787-184
; Sequence 184, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 184
; LENGTH: 744
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-184

Query Match          96.9%; Score 498; DB 10; Length 744;
Best Local Similarity 98.1%; Pred. No. 2e-33;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
    |||||
Db 346 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 405
    |||||

Qy 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104
    |||||
Db 406 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 449
    |||||

RESULT 4
US-10-472-928-32
; Sequence 32, Application US/10472928
; Publication No. US20050020813A1
```

```
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 32
; LENGTH: 744
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: pneumococcal surface protein A (pspA)
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Feature of note: WYY motif
; OTHER INFORMATION: Similar to strain R6 sequence 15902165 (e-179)
US-10-472-928-32

Query Match          96.9%; Score 498; DB 17; Length 744;
Best Local Similarity 98.1%; Pred. No. 2e-33;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 60
    |||||
Db 346 LAKQTELEKLLSLDPEGKTQDELKAEAEALDKKADELPNKVADLEKEISNLEILLG 405
    |||||

Qy 61 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 104
    |||||
Db 406 GADSEDDTAALPNKLTAKAELEKTQKELDAALNELGPDGDEEE 449
    |||||

RESULT 5
US-09-765-272-160
; Sequence 160, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-765-272-160

Query Match          96.1%; Score 494; DB 9; Length 641;
Best Local Similarity 97.1%; Pred. No. 3.6e-33;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
    |||||
Db 243 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 302
    |||||

QY 61 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEEE 104
    |||||
Db 303 GADXEDDTAALQNKLTATKKAELKTKQKELDAALNELGPDGDEEE 346
    |||||

RESULT 6
US-10-299-636-62
; Sequence 62, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; PRIOR FILING DATE: 2002-11-19
; PRIOR FILING DATE: 1996-09-16
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa at position 2 is unknown
US-10-299-636-62

Query Match          95.3%; Score 490; DB 15; Length 213;
Best Local Similarity 95.2%; Pred. No. 2.1e-33;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
    |||||
Db 59 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 118
    |||||

QY 61 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEEE 104
    |||||
Db 119 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEEE 162
    |||||

RESULT 7
US-10-299-636-59
; Sequence 59, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
```

```
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-299-636-59

Query Match          94.6%; Score 486; DB 15; Length 197;
Best Local Similarity 95.2%; Pred. No. 4.2e-33;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
    |||||
Db 22 LAQKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 81
    |||||

QY 61 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEEE 104
    |||||
Db 82 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEEE 125
    |||||

RESULT 8
US-10-674-755-18
; Sequence 18, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-18

Query Match          94.2%; Score 484; DB 15; Length 102;
Best Local Similarity 98.1%; Pred. No. 2.8e-33;
Matches 102; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 60
    |||||
Db 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEAEADKKADELPNKVADLEKEISNLEILLG 58
    |||||

QY 61 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEEE 104
    |||||
Db 59 GADSEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEEE 102
    |||||

RESULT 9
US-10-299-636-67
; Sequence 67, Application US/10299636
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
```

```

; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; OTHER INFORMATION: Xaa at position 129 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (131)
; OTHER INFORMATION: Xaa at position 131 is unknown
; US-10-299-636-67

Query Match          90.5%; Score 465; DB 15; Length 233;
Best Local Similarity 92.3%; Pred. No. 2.9e-31;
Matches 96; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELDKAEAEAEELDKKADLPNKVADLEKISNLEILLG 60
||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
53 LAKPTGLEKLLDSLDPEGKTQDELDKAEAEAEELDKKADLPNKVADLEKISNLEILLG 112
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 104
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPDGDEEE 156
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-10-674-755-19
; Sequence 19, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-10-674-755-19

Query Match          73.5%; Score 378; DB 15; Length 80;
Best Local Similarity 96.2%; Pred. No. 1.7e-24;
Matches 77; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 25 LDKAEAEALDKKADLPNKVADLEKISNLEILLGGADSEDDTAALPNKLTAKAELEK 84
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 LDKAEAEALDKKADLPNKVADLEKISNLEILLGGADSEDDTAALPNKLTAKAELEK 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 85 TOKELDAALNELGPDGDEEE 104
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



QY 1 LAKKOTELEKLDSLDPEGKTQDELCKEAEFAELDKKADELPNKVADLEKEISNLEILG 60

QY 6I GADS---ED-DIALEFNALAINAELEKIQREDDAALNEDGFDDEE 1V4

Search completed: June 21, 2005, 11:18:38  
Job time : 67.4327 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:53:39 ; Search time 10.4 Seconds  
(without alignments)  
962.168 Million cell updates/sec

Title: US-10-674-755-20

Perfect score: 514

Sequence: 1 LAKKQTELEKLLDLSLDPGK.....TQKELDAALNELGPDGDEEE 104

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 498   | 96.9        | 744    | 2 P95013 | pneumococcal surfa |
| 2          | 128   | 24.9        | 619    | 2 A97887 | surface protein ps |
| 3          | 128   | 24.9        | 619    | 2 A41971 | surface protein ps |
| 4          | 109.5 | 21.3        | 388    | 2 A46173 | Mrp4 protein - Str |
| 5          | 108   | 21.0        | 1837   | 2 T41023 | probable nuclear p |
| 6          | 107   | 20.8        | 385    | 2 T20410 | hypothetical prote |
| 7          | 106.5 | 20.7        | 1509   | 1 A27224 | myosin heavy chain |
| 8          | 104.5 | 20.3        | 388    | 2 S52536 | fcrA 15 protein -  |
| 9          | 104.5 | 20.3        | 924    | 2 S06117 | myosin heavy chain |
| 10         | 104.5 | 20.3        | 2007   | 1 B43402 | myosin heavy chain |
| 11         | 104   | 20.2        | 281    | 2 F75216 | hypothetical prote |
| 12         | 103   | 20.0        | 1964   | 2 A59282 | nonmuscle myosin I |
| 13         | 102.5 | 19.9        | 405    | 2 A33939 | Fc gamma (IgG) rec |
| 14         | 102.5 | 19.9        | 2139   | 2 T18296 | myosin heavy chain |
| 15         | 101.5 | 19.7        | 1937   | 2 I38055 | myosin heavy chain |
| 16         | 101   | 19.6        | 1175   | 2 C35815 | myosin heavy chain |
| 17         | 101   | 19.6        | 1175   | 2 D35815 | myosin heavy chain |
| 18         | 101   | 19.6        | 1201   | 2 B35815 | myosin heavy chain |
| 19         | 101   | 19.6        | 1201   | 2 B35815 | myosin heavy chain |
| 20         | 101   | 19.6        | 2385   | 2 A32491 | myosin heavy chain |
| 21         | 101   | 19.6        | 2411   | 2 B32491 | myosin heavy chain |
| 22         | 100.5 | 19.6        | 629    | 2 T44607 | hypothetical prote |
| 23         | 100.5 | 19.6        | 1938   | 1 A40997 | myosin heavy chain |
| 24         | 100   | 19.5        | 259    | 2 D60110 | repetitive protein |
| 25         | 100   | 19.5        | 318    | 2 T49167 | hypothetical prote |
| 26         | 99.5  | 19.4        | 397    | 2 H86754 | prophage pi2 prote |
| 27         | 99.5  | 19.4        | 1992   | 1 S02771 | myosin heavy chain |
| 28         | 99    | 19.3        | 516    | 2 B84709 | hypothetical prote |
| 29         | 98.5  | 19.2        | 527    | 2 S33068 | myosin heavy chain |

30 98.5 19.2 1940 2 A59287 myosin heavy chain  
31 98.5 19.2 1976 2 A59252 myosin heavy chain  
32 98 19.1 248 2 T26412 hypothetical prote  
33 98 19.1 1169 2 A64505 p115 homolog - Met  
34 98 19.1 1939 2 T18372 repeat organellar  
35 97.5 19.0 284 2 A45488 body-wall muscle t  
36 97.5 19.0 387 2 S57834 fcrA protein precu  
37 97.5 19.0 1119 2 T14321 nuclear matrix con  
38 97.5 19.0 1475 2 T33318 hypothetical prote  
39 97 18.9 518 2 G84488 En/Spm-like transp  
40 97 18.9 587 2 JC1419 Fc gamma (IgG) rec  
41 97 18.9 1063 2 T18255 cytoskeleton assem  
42 97 18.9 1179 2 P71190 probable chromosom  
43 96.5 18.8 281 2 A34787 tropomyosin 1 alph  
44 96.5 18.8 284 1 TMRBA tropomyosin alpha  
45 96.5 18.8 284 2 A39816 tropomyosin 2, fib

#### ALIGNMENTS

##### RESULT 1

F95013

pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C;Accession: F95013

R;Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95013

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-744 <KUR>

A;Cross-references: UNIPROT:Q97T39; GB:AE005672; PIDN:AAK74303.1; PID:gl14971584; GSPDB:G

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP0117

Query Match 96.9%; Score 498; DB 2; Length 744;  
Best Local Similarity 98.1%; Pred. No. 2.2e-26;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKTQDELKAEAEAEALDKADELPKNKADLEKEISNLEILLG 60

|||||

DB 346 LAKKQTELEKLLDLSLDPGKTQDELKAEAEAEALDKADELPKNKADLEKEISNLEILLG 405

|||||

QY 61 GADSEDDTAALPNKLATKKALEKTKOKELDAALNELGPDGDEEE 104

|||||

DB 406 GADSEDDTAALQNKATKKALEKTKOKELDAALNELGPDGDEEE 449

|||||

##### RESULT 2

A97887

surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C;Accession: A97887

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Es

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5703-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: A97887

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-619 <KUR>

A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:gl  
C;Genetics:  
A;Gene: pspA

Query Match 24.9%; Score 128; DB 2; Length 619;  
Best Local Similarity 33.6%; Pred. No. 0.14;  
Matches 38; Conservative 19; Mismatches 34; Indels 22; Gaps 3;

Qy 4 KQTEKLKLSLDPEGKTQDELKDE-----AREAEIDKKADELPNKVAD 47  
Db 208 KIALENQVHLEQELKEIDSESEDYAKGFRAPLOSKLDKAKKL-SKLEELSDKIDE 266

Qy 48 LEKEISNLEILLGASDDT-----AALPNKLATKKALEKTKQKELDAALNE 95  
Db 267 LDABIAKLELDQKAAEENNVEDYFKEGLEKTIAAKKALEKTEADLKAVNE 319

RESULT 3  
A41971  
surface protein pspA precursor - Streptococcus pneumoniae  
N;Alternate names: pneumococcal surface protein A  
C;Species: Streptococcus pneumoniae  
C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
A;Accession: A41971; A60282; A3134  
R;Yother, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A;Title: Structural properties and evolutionary relationships of PspA, a surface protein  
A;Reference number: A41971; MUID:92105030; PMID:1729249  
A;Accession: A41971  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-619 <YOT>  
A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:g153840; PIDN:AAA2701  
A;Note: sequence extracted from NCBI backbone (NCBIN:75635, NCBIP:75636)  
R;Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yother, J.; Briles, D.E.  
Infect. Immun. 59, 1285-1289, 1991  
A;Title: A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective ability  
A;Reference number: A60282; MUID:91169598; PMID:2004810  
A;Accession: A60282  
A;Molecule type: protein  
A;Residues: 32-76 <TAL>  
A;Experimental source: strain JY2008  
C;Genetics:  
A;Gene: pspA  
F;1-31/Domain: signal sequence #status predicted <SIG>  
F;32-619/Product: surface protein pspA #status predicted <MAT>  
F;411-430/Domain: cpl repeat homology <CP01>  
F;431-450/Domain: cpl repeat homology <CP02>  
F;451-470/Domain: cpl repeat homology <CP03>  
F;471-490/Domain: cpl repeat homology <CP04>  
F;491-510/Domain: cpl repeat homology <CP05>  
F;511-530/Domain: cpl repeat homology <CP06>  
F;531-550/Domain: cpl repeat homology <CP07>  
F;551-570/Domain: cpl repeat homology <CP08>  
F;571-591/Domain: cpl repeat homology <CP09>  
F;592-611/Domain: cpl repeat homology <CP10>

Query Match 24.9%; Score 128; DB 2; Length 619;  
Best Local Similarity 33.6%; Pred. No. 0.14; Mismatches 34; Indels 22; Gaps 3;

Qy 4 KQTEKLKLSLDPEGKTQDELKDE-----AREAEIDKKADELPNKVAD 47  
Db 208 KIALENQVHLEQELKEIDSESEDYAKGFRAPLOSKLDKAKKL-SKLEELSDKIDE 266

Qy 48 LEKEISNLEILLGASDDT-----AALPNKLATKKALEKTKQKELDAALNE 95  
Db 267 LDABIAKLELDQKAAEENNVEDYFKEGLEKTIAAKKALEKTEADLKAVNE 319

RESULT 4  
A46173  
MrpA protein - Streptococcus sp. (group A)

C;Species: Streptococcus sp.  
C;Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 30-May-1997  
C;Accession: A46173  
R;O'Toole, P.; Stenberg, L.; Rissler, M.; Lindahl, G.  
Proc. Natl. Acad. Sci. U.S.A. 89, 8661-8665, 1992  
A;Title: Two major classes in the M protein family in group A streptococci.  
A;Reference number: A46173; MUID:92409576; PMID:1528877  
A;Contents: group A  
A;Accession: A46173  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-388 <OIT>  
A;Note: sequence extracted from NCBI backbone (NCBIN:114063, NCBIP:114064)  
C;Superfamily: M5 protein

Query Match 21.3%; Score 109.5; DB 2; Length 388;  
Best Local Similarity 33.0%; Pred. No. 1.5;  
Matches 36; Conservative 25; Mismatches 41; Indels 7; Gaps 4;

Qy 1 LAKKQTE---LEKLLDSLDPE-GKTQDELQ-KEAEAEALDKKADELPNKVADLEKEISNL 55  
Db 182 IAKQSEATLENLGSAKRELTELQAKLDTATAEKAKLBSQVTTLENLGSAKRELTDL 241

Qy 56 EILLGGADSEDDTAALPNKLATKKALEKTKQKELDAALNEILGPDGDEEE 104  
Db 242 QAKLDAANAEE--KLQSQATLEKQLEATKKELDLQAKLAATNQEKE 288

RESULT 5  
T41023  
probable nuclear pore complex-associated protein - fission yeast (Schizosaccharomyces po  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T41023  
R;Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, June 1998  
A;Reference number: Z21965  
A;Accession: T41023  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-1837 <MUR>  
A;Cross-references: UNIPROT:O74424; EMBL:AL023860; PIDN:CAA19588.1; GSPDB:GN00068; SPDB:5  
A;Experimental source: strain 972h-; cosmid c162  
C;Genetics:  
A;Gene: SPDB:SPCC162.08C  
A;Map position: 3

Query Match 21.0%; Score 108; DB 2; Length 1837;  
Best Local Similarity 31.1%; Pred. No. 9;  
Matches 37; Conservative 21; Mismatches 35; Indels 26; Gaps 5;

Qy 1 LAKKQTELEKL-LDSLDPEGKTQ-DELDKEAEAEALDKKADELPNKVADLEKEISNL--- 55  
Db 1500 LADSKNEHLQSEAVDADGKTISNLEKEIHELRSDEG--LVQQVQNLUSAEALAUHEH 1557

Qy 56 EILLGGADSEDDTAALPNKLATKK-----AELEKTKQKELDAALNE 95  
Db 1558 SPTQGSLENADELARLSQLESTKQYYEKEKETELLAARSELVAEREKTKKEELNQLNE 1616

RESULT 6  
T20410  
hypothetical protein E02A10.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T20410  
R;Thomas, K.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z19271  
A;Accession: T20410  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-385 <WIL>

A;Accession: S52536  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-388 <KAT>  
A;Cross-references: UNIPROT:Q53474; GB:S75411; NID:g914107; PIDN:AAB33261.1; PID:g914109  
C;Superfamily: M5 protein

Query Match 20.3%; Score 104.5; DB 2; Length 388;  
Best Local Similarity 33.0%; Pred. No. 3.2;  
Matches 36; Conservative 23; Mismatches 43; Indels 7; Gaps 4;

Qy 1 LAKKOTE---LEKLDSLDPG-GKTQDELDDKEAEAEELDKADELPKNKVADLEKEISNL 55  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
182 IAKQSEAAITLENLGSARELTDLQAKUTATAKAKLESQVTTLENLIGSAKRELTDL 241  
Qy 56 EILGGADSEDDTAALPNKLYATKKAELEKTKQKELDAAALNELGPDGDDEE 104  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
242 QAKLDANAEE--KAKLSQAALAEKLEATKKELADLQAKLAATNQEK 288

RESULT 9  
S06117  
myosin heavy chain, nonmuscle (clone lambda-FWHC) - chicken (fragment)  
C;Species: Gallus gallus (Chicken)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: S06117  
R;Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Masaki, T.  
Eur. J. Biochem. 184, 611-616, 1989  
A;Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in  
S.

A;Reference number: S06116; MUID:90032648; PMID:2806244  
A;Accession: S06117  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-924 <KAT>  
A;Cross-references: UNIPROT:Q02015; GB:X17590  
A;Note: This translation is not annotated in GenBank entry GGMHCF, release 114  
C;Superfamily: myosin heavy chain; myosin motor domain homology  
F;1-303/Domain: myosin motor domain homology (fragment) <MMOT>

Query Match 20.3%; Score 104.5; DB 2; Length 924;  
Best Local Similarity 30.3%; Pred. No. 7.7;  
Matches 33; Conservative 26; Mismatches 43; Indels 7; Gaps 3;

Qy 1 LAKKQTELEKLDSLD----PEGKTQDELDDKEAEAEELDKADELPKNKVADLEKEISNL 56  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
557 LAKLNKGEMMITDLEERLKKEETKQELEK--AKRKLDGETTDLQDAELQAIQIEELK 614

Qy 57 ILGGADSEDDTA-ALPNKLYATKKAELEKTKQKELDAAALNELGPDGDDEE 104  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
615 IQLAKKEBELQALARGDEAVQKNALKVIRELQAIQAELEDLESEK 663

RESULT 10  
B43402  
myosin heavy chain-B, neuronal - chicken  
N;Contains: myosin ATPase (EC 3.6.4.1)  
C;Species: Gallus gallus (Chicken)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: B43402; A43402  
R;Takahashi, M.; Kawamoto, S.; Adelstein, R.S.  
J. Biol. Chem. 267, 17864-17871, 1992  
A;Title: Evidence for inserted sequences in the head region of nonmuscle myosin specific  
Yosin.

A;Reference number: A43402; MUID:92388144; PMID:1355479  
A;Accession: B43402  
A;Molecule type: mRNA  
A;Residues: 1-2007 <TAK>  
A;Cross-references: UNIPROT:Q02015; GB:M93676; NID:g212448; PIDN:AAA48988.1; PID:g212452  
A;Note: The sequence of residues 212-221 and 632-652 and the corresponding nucleotide se  
A;Accession: A43402  
A;Molecule type: mRNA  
A;Residues: 1-211;222-631;653-2007 <TA2>



C:Superfamily: myosin heavy chain; myosin motor domain homology  
 F;91-780/Domain: myosin motor domain homology <MMOT>

Query Match 19.9%; Score 102.5; DB 2; Length 2139;  
 Best Local Similarity 32.4%; Pred. No. 25;  
 Matches 35; Conservative 16; Mismatches 36; Indels 21; Gaps 4;

QY 1 LAKKQTELEKLLSLDPGKGTQDELKAEAEALDKADLPNKNVADLEKEISNLE---I 57  
 Db 1507 VAKLNTQITKL-----TRDQSAEELNLSRKADKKKIKISELEQVNELESRPV 1557

QY 58 LLGGADS-----EDDTAALPNKLTAKKA-----ELEKTKDELDAALNEL 96  
 Db 1558 GTGNADENEIKIRDAIADLNKALEMKGVQNNQATNKNELKAKNDNL 1605

RESULT 15  
 I38055  
 myosin heavy chain, perinatal skeletal muscle - human  
 N;Contains: myosin ATPase (EC 3.6.4.1)  
 C;Species: Homo sapiens (man)  
 C;Date: 17-May-1996 #sequence revision 17-May-1996 #text change 09-Jul-2004  
 C;Accession: I38055; JH0154; S12459; S09332; A30220; S49478  
 R;Jullian, E.H.; Kelly, A.M.; Pompidou, A.J.; Hoffman, R.; Schiaffino, S.; Stedman, H.H.  
 Eur. J. Biochem. 230, 1001-1006, 1995  
 A;Title: Characterization of a human perinatal myosin heavy-chain transcript.  
 A;Reference number: I38055; MUID:95324556; PMID:7601129  
 A;Accession: I38055  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-1937 <RES>  
 A;Cross-references: UNIPROT:P13535; EMBL:X51592; NID:9558668; PIDN:CAA86293.1; PID:9558668  
 R;Karsch-Mizrachi, I.; Feghali, R.; Shows, T.B.; Leinwand, L.A.  
 Gene 89, 289-294, 1990  
 A;Title: Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.  
 A;Reference number: JH0154; MUID:90323631; PMID:2373371  
 A;Accession: JH0154  
 A;Molecule type: mRNA  
 A;Residues: 1-14, 'A', 16-859 <KAR>  
 A;Cross-references: GB:Y00821  
 A;Experimental source: skeletal muscle  
 R;Bober, E.  
 submitted to the EMBL Data Library, January 1989  
 A;Reference number: S12458  
 A;Accession: S12459  
 A;Molecule type: mRNA  
 A;Residues: 502-1071, 'N', 1073-1250, 'DGG', 1253-1376, 'NT', 1379-1913, 'D', 1915-1937 <BOB>  
 A;Cross-references: EMBL:X51592; NID:929465; PIDN:CAA35941.1; PID:929466  
 A;Experimental source: clone gCMC-F  
 R;Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.  
 Eur. J. Biochem. 189, 55-65, 1990  
 A;Title: Identification of three developmentally controlled isoforms of human myosin heavy chain  
 A;Reference number: S09331; MUID:90235862; PMID:1691980  
 A;Accession: S09332  
 A;Molecule type: mRNA  
 A;Residues: 502-547, 'X', 549-617, 'X', 619-687, 'X', 689-757, 'X', 759-827, 'X', 829-897, 'X', 899-1376, 'NT', 1379-1386, 'X', 1388-1456, 'X', 1458-1526, 'X', 1528-1596, 'X', 1598-1666, 'X', 1668-1708, 'X', 1710-1717, 'X', 1719-1726, 'X', 1728-1735, 'X', 1737-1744, 'X', 1746-1753, 'X', 1755-1762, 'X', 1764-1771, 'X', 1773-1780, 'X', 1782-1789, 'X', 1791-1797, 'X', 1799-1806, 'X', 1808-1815, 'X', 1817-1824, 'X', 1826-1833, 'X', 1835-1842, 'X', 1844-1851, 'X', 1853-1860, 'X', 1862-1869, 'X', 1871-1878, 'X', 1880-1887, 'X', 1889-1896, 'X', 1898-1905, 'X', 1907-1914, 'X', 1916-1923, 'X', 1925-1932, 'X', 1934-1941, 'X', 1943-1950, 'X', 1952-1959, 'X', 1961-1968, 'X', 1970-1977, 'X', 1979-1986, 'X', 1988-1995, 'X', 1997-2004, 'X', 2006-2013, 'X', 2015-2022, 'X', 2024-2031, 'X', 2033-2040, 'X', 2042-2049, 'X', 2051-2058, 'X', 2060-2067, 'X', 2069-2076, 'X', 2078-2085, 'X', 2087-2094, 'X', 2096-2103, 'X', 2105-2112, 'X', 2114-2121, 'X', 2123-2130, 'X', 2132-2139, 'X', 2141-2148, 'X', 2150-2157, 'X', 2159-2166, 'X', 2168-2175, 'X', 2177-2184, 'X', 2186-2193, 'X', 2195-2202, 'X', 2204-2211, 'X', 2213-2220, 'X', 2222-2229, 'X', 2231-2238, 'X', 2240-2247, 'X', 2249-2256, 'X', 2258-2265, 'X', 2267-2274, 'X', 2276-2283, 'X', 2285-2292, 'X', 2294-2301, 'X', 2303-2310, 'X', 2312-2319, 'X', 2321-2328, 'X', 2330-2337, 'X', 2339-2346, 'X', 2348-2355, 'X', 2357-2364, 'X', 2366-2373, 'X', 2375-2382, 'X', 2384-2391, 'X', 2393-2400, 'X', 2402-2409, 'X', 2411-2418, 'X', 2420-2427, 'X', 2429-2436, 'X', 2438-2445, 'X', 2447-2454, 'X', 2456-2463, 'X', 2465-2472, 'X', 2474-2481, 'X', 2483-2490, 'X', 2492-2499, 'X', 2501-2508, 'X', 2510-2517, 'X', 2519-2526, 'X', 2528-2535, 'X', 2537-2544, 'X', 2546-2553, 'X', 2555-2562, 'X', 2564-2571, 'X', 2573-2580, 'X', 2582-2589, 'X', 2591-2598, 'X', 2600-2607, 'X', 2609-2616, 'X', 2618-2625, 'X', 2627-2634, 'X', 2636-2643, 'X', 2645-2652, 'X', 2654-2661, 'X', 2663-2670, 'X', 2672-2679, 'X', 2681-2688, 'X', 2690-2697, 'X', 2699-2706, 'X', 2708-2715, 'X', 2717-2724, 'X', 2726-2733, 'X', 2735-2742, 'X', 2744-2751, 'X', 2753-2760, 'X', 2762-2769, 'X', 2771-2778, 'X', 2780-2787, 'X', 2789-2796, 'X', 2798-2805, 'X', 2807-2814, 'X', 2816-2823, 'X', 2825-2832, 'X', 2834-2841, 'X', 2843-2850, 'X', 2852-2859, 'X', 2861-2868, 'X', 2870-2877, 'X', 2879-2886, 'X', 2888-2895, 'X', 2897-2904, 'X', 2906-2913, 'X', 2915-2922, 'X', 2924-2931, 'X', 2933-2940, 'X', 2942-2949, 'X', 2951-2958, 'X', 2960-2967, 'X', 2969-2976, 'X', 2978-2985, 'X', 2987-2994, 'X', 2996-3003, 'X', 3005-3012, 'X', 3014-3021, 'X', 3023-3030, 'X', 3032-3039, 'X', 3041-3048, 'X', 3050-3057, 'X', 3059-3066, 'X', 3068-3075, 'X', 3077-3084, 'X', 3086-3093, 'X', 3095-3102, 'X', 3104-3111, 'X', 3113-3120, 'X', 3122-3129, 'X', 3131-3138, 'X', 3140-3147, 'X', 3149-3156, 'X', 3158-3165, 'X', 3167-3174, 'X', 3176-3183, 'X', 3185-3192, 'X', 3194-3201, 'X', 3203-3210, 'X', 3212-3219, 'X', 3221-3228, 'X', 3230-3237, 'X', 3239-3246, 'X', 3248-3255, 'X', 3257-3264, 'X', 3266-3273, 'X', 3275-3282, 'X', 3284-3291, 'X', 3293-3300, 'X', 3302-3309, 'X', 3311-3318, 'X', 3320-3327, 'X', 3329-3336, 'X', 3338-3345, 'X', 3347-3354, 'X', 3356-3363, 'X', 3365-3372, 'X', 3374-3381, 'X', 3383-3390, 'X', 3392-3399, 'X', 3401-3408, 'X', 3410-3417, 'X', 3419-3426, 'X', 3428-3435, 'X', 3437-3444, 'X', 3446-3453, 'X', 3455-3462, 'X', 3464-3471, 'X', 3473-3480, 'X', 3482-3489, 'X', 3491-3498, 'X', 3499-3506, 'X', 3508-3515, 'X', 3517-3524, 'X', 3526-3533, 'X', 3535-3542, 'X', 3544-3551, 'X', 3553-3560, 'X', 3562-3569, 'X', 3571-3578, 'X', 3580-3587, 'X', 3589-3596, 'X', 3598-3605, 'X', 3607-3614, 'X', 3616-3623, 'X', 3625-3632, 'X', 3634-3641, 'X', 3643-3650, 'X', 3652-3659, 'X', 3661-3668, 'X', 3670-3677, 'X', 3679-3686, 'X', 3688-3695, 'X', 3697-3704, 'X', 3706-3713, 'X', 3715-3722, 'X', 3724-3731, 'X', 3733-3740, 'X', 3742-3749, 'X', 3751-3758, 'X', 3760-3767, 'X', 3769-3776, 'X', 3778-3785, 'X', 3787-3794, 'X', 3796-3803, 'X', 3805-3812, 'X', 3814-3821, 'X', 3823-3830, 'X', 3832-3839, 'X', 3841-3848, 'X', 3850-3857, 'X', 3859-3866, 'X', 3868-3875, 'X', 3877-3884, 'X', 3886-3893, 'X', 3895-3902, 'X', 3904-3911, 'X', 3913-3920, 'X', 3922-3929, 'X', 3931-3938, 'X', 3940-3947, 'X', 3949-3956, 'X', 3958-3965, 'X', 3967-3974, 'X', 3976-3983, 'X', 3985-3992, 'X', 3994-4001, 'X', 4003-4010, 'X', 4012-4019, 'X', 4021-4028, 'X', 4030-4037, 'X', 4039-4046, 'X', 4048-4055, 'X', 4057-4064, 'X', 4066-4073, 'X', 4075-4082, 'X', 4084-4091, 'X', 4093-4100, 'X', 4102-4109, 'X', 4111-4118, 'X', 4120-4127, 'X', 4129-4136, 'X', 4138-4145, 'X', 4147-4154, 'X', 4156-4163, 'X', 4165-4172, 'X', 4174-4181, 'X', 4183-4190, 'X', 4192-4199, 'X', 4201-4208, 'X', 4210-4217, 'X', 4219-4226, 'X', 4228-4235, 'X', 4237-4244, 'X', 4246-4253, 'X', 4255-4262, 'X', 4264-4271, 'X', 4273-4280, 'X', 4282-4289, 'X', 4291-4298, 'X', 4299-4306, 'X', 4308-4315, 'X', 4317-4324, 'X', 4326-4333, 'X', 4335-4342, 'X', 4344-4351, 'X', 4353-4360, 'X', 4362-4369, 'X', 4371-4378, 'X', 4380-4387, 'X', 4389-4396, 'X', 4398-4405, 'X', 4407-4414, 'X', 4416-4423, 'X', 4425-4432, 'X', 4434-4441, 'X', 4443-4450, 'X', 4452-4459, 'X', 4461-4468, 'X', 4470-4477, 'X', 4479-4486, 'X', 4488-4495, 'X', 4497-4504, 'X', 4506-4513, 'X', 4515-4522, 'X', 4524-4531, 'X', 4533-4540, 'X', 4542-4549, 'X', 4551-4558, 'X', 4560-4567, 'X', 4569-4576, 'X', 4578-4585, 'X', 4587-4594, 'X', 4596-4603, 'X', 4605-4612, 'X', 4614-4621, 'X', 4623-4630, 'X', 4632-4639, 'X', 4641-4648, 'X', 4650-4657, 'X', 4659-4666, 'X', 4668-4675, 'X', 4677-4684, 'X', 4686-4693, 'X', 4695-4702, 'X', 4704-4711, 'X', 4713-4720, 'X', 4722-4729, 'X', 4731-4738, 'X', 4740-4747, 'X', 4749-4756, 'X', 4758-4765, 'X', 4767-4774, 'X', 4776-4783, 'X', 4785-4792, 'X', 4794-4801, 'X', 4803-4810, 'X', 4812-4819, 'X', 4821-4828, 'X', 4830-4837, 'X', 4839-4846, 'X', 4848-4855, 'X', 4857-4864, 'X', 4866-4873, 'X', 4875-4882, 'X', 4884-4891, 'X', 4893-4900, 'X', 4902-4909, 'X', 4911-4918, 'X', 4920-4927, 'X', 4929-4936, 'X', 4938-4945, 'X', 4947-4954, 'X', 4956-4963, 'X', 4965-4972, 'X', 4974-4981, 'X', 4983-4990, 'X', 4992-4999, 'X', 5001-5008, 'X', 5010-5017, 'X', 5019-5026, 'X', 5028-5035, 'X', 5037-5044, 'X', 5046-5053, 'X', 5055-5062, 'X', 5064-5071, 'X', 5073-5080, 'X', 5082-5089, 'X', 5091-5098, 'X', 5099-5106, 'X', 5108-5115, 'X', 5117-5124, 'X', 5126-5133, 'X', 5135-5142, 'X', 5144-5151, 'X', 5153-5160, 'X', 5162-5169, 'X', 5171-5178, 'X', 5180-5187, 'X', 5189-5196, 'X', 5198-5205, 'X', 5207-5214, 'X', 5216-5223, 'X', 5225-5232, 'X', 5234-5241, 'X', 5243-5250, 'X', 5252-5259, 'X', 5261-5268, 'X', 5270-5277, 'X', 5279-5286, 'X', 5288-5295, 'X', 5297-5304, 'X', 5306-5313, 'X', 5315-5322, 'X', 5324-5331, 'X', 5333-5340, 'X', 5342-5349, 'X', 5351-5358, 'X', 5360-5367, 'X', 5369-5376, 'X', 5378-5385, 'X', 5387-5394, 'X', 5396-5403, 'X', 5405-5412, 'X', 5414-5421, 'X', 5423-5430, 'X', 5432-5439, 'X', 5441-5448, 'X', 5450-5457, 'X', 5459-5466, 'X', 5468-5475, 'X', 5477-5484, 'X', 5486-5493, 'X', 5495-5502, 'X', 5504-5511, 'X', 5513-5520, 'X', 5522-5529, 'X', 5531-5538, 'X', 5540-5547, 'X', 5549-5556, 'X', 5558-5565, 'X', 5567-5574, 'X', 5576-5583, 'X', 5585-5592, 'X', 5594-5601, 'X', 5603-5610, 'X', 5612-5619, 'X', 5621-5628, 'X', 5630-5637, 'X', 5639-5646, 'X', 5648-5655, 'X', 5657-5664, 'X', 5666-5673, 'X', 5675-5682, 'X', 5684-5691, 'X', 5693-5700, 'X', 5702-5709, 'X', 5711-5718, 'X', 5720-5727, 'X', 5729-5736, 'X', 5738-5745, 'X', 5747-5754, 'X', 5756-5763, 'X', 5765-5772, 'X', 5774-5781, 'X', 5783-5790, 'X', 5792-5799, 'X', 5799-5806, 'X', 5808-5815, 'X', 5817-5824, 'X', 5826-5833, 'X', 5835-5842, 'X', 5844-5851, 'X', 5853-5860, 'X', 5862-5869, 'X', 5871-5878, 'X', 5880-5887, 'X', 5889-5896, 'X', 5900-5907, 'X', 5909-5916, 'X', 5918-5925, 'X', 5927-5934, 'X', 5936-5943, 'X', 5945-5952, 'X', 5954-5961, 'X', 5963-5970, 'X', 5972-5979, 'X', 5981-5988, 'X', 5990-5997, 'X', 5999-6006, 'X', 6008-6015, 'X', 6017-6024, 'X', 6026-6033, 'X', 6035-6042, 'X', 6044-6051, 'X', 6053-6060, 'X', 6062-6069, 'X', 6071-6078, 'X', 6080-6087, 'X', 6089-6096, 'X', 6098-6105, 'X', 6107-6114, 'X', 6116-6123, 'X', 6125-6132, 'X', 6134-6141, 'X', 6143-6150, 'X', 6152-6159, 'X', 6161-6168, 'X', 6170-6177, 'X', 6179-6186, 'X', 6188-6195, 'X', 6197-6204, 'X', 6206-6213, 'X', 6215-6222, 'X', 6224-6231, 'X', 6233-6240, 'X', 6242-6249, 'X', 6251-6258, 'X', 6260-6267, 'X', 6269-6276, 'X', 6278-6285, 'X', 6287-6294, 'X', 6296-6303, 'X', 6305-6312, 'X', 6314-6321, 'X', 6323-6330, 'X', 6332-6339, 'X', 6341-6348, 'X', 6350-6357, 'X', 6359-6366, 'X', 6368-6375, 'X', 6377-6384, 'X', 6386-6393, 'X', 6395-6402, 'X', 6404-6411, 'X', 6413-6420, 'X', 6422-6429, 'X', 6431-6438, 'X', 6440-6447, 'X', 6449-6456, 'X', 6458-6465, 'X', 6467-6474, 'X', 6476-6483, 'X', 6485-6492, 'X', 6494-6501, 'X', 6503-6510, 'X', 6512-6519, 'X', 6521-6528, 'X', 6530-6537, 'X', 6539-6546, 'X', 6548-6555, 'X', 6557-6564, 'X', 6566-6573, 'X', 6575-6582, 'X', 6584-6591, 'X', 6593-6600, 'X', 6602-6609, 'X', 6611-6618, 'X', 6620-6627, 'X', 6629-6636, 'X', 6638-6645, 'X', 6647-6654, 'X', 6656-6663, 'X', 6665-6672, 'X', 6674-6681, 'X', 6683-6690, 'X', 6692-6699, 'X', 6699-6706, 'X', 6708-6715, 'X', 6717-6724, 'X', 6726-6733, 'X', 6735-6742, 'X', 6744-6751, 'X', 6753-6760, 'X', 6762-6769, 'X', 6771-6778, 'X', 6779-6786, 'X', 6788-6795, 'X', 6797-6804, 'X', 6806-6813, 'X', 6815-6822, 'X', 6824-6831, 'X', 6833-6840, 'X', 6842-6849, 'X', 6851-6858, 'X', 6860-6867, 'X', 6869-6876, 'X', 6878-6885, 'X', 6887-6894, 'X', 6896-6903, 'X', 6905-6912, 'X', 6914-6921, 'X', 6923-6930, 'X', 6932-6939, 'X', 6941-6948, 'X', 6950-6957, 'X', 6959-6966, 'X', 6968-6975, 'X', 6977-6984, 'X', 6986-6993, 'X', 6995-7002, 'X', 7004-7011, 'X', 7013-7020, 'X', 7022-7029, 'X', 7031-7038, 'X', 7040-7047, 'X', 7049-7056, 'X', 7058-7065, 'X', 7067-7074, 'X', 7076-7083, 'X', 7085-7092, 'X', 7094-7101, 'X', 7103-7110, 'X', 7112-7119, 'X', 7121-7128, 'X', 7130-7137, 'X', 7139-7146, 'X', 7148-7155, 'X', 7157-7164, 'X', 7166-7173, 'X', 7175-7182, 'X', 7184-7191, 'X', 7193-7200, 'X', 7202-7209, 'X', 7211-7218, 'X', 7220-7227, 'X', 7229-7236, 'X', 7238-7245, 'X', 7247-7254, 'X', 7256-7263, 'X', 7265-7272, 'X', 7274-7281, 'X', 7283-7290, 'X', 7292-7299, 'X', 7301-7308, 'X', 7310-7317, 'X', 7319-7326, 'X', 7328-7335, 'X', 7337-7344, 'X', 7346-7353, 'X', 7355-7362, 'X', 7364-7371, 'X', 7373-7380, 'X', 7382-7389, 'X', 7391-7398, 'X', 7399-7406, 'X', 7408-7415, 'X', 7417-7424, 'X', 7426-7433, 'X', 7435-7442, 'X', 7444-7451, 'X', 7453-7460, 'X', 7462-7469, 'X', 7471-7478, 'X', 7480-7487, 'X', 7489-7496, 'X', 7498-7505, 'X', 7507-7514, 'X', 7516-7523, 'X', 7525-7532, 'X', 7534-7541, 'X', 7543-7550, 'X', 7552-7559, 'X', 7561-7568, 'X', 7570-7577, 'X', 7579-7586, 'X', 7588-7595, 'X', 7597-7604, 'X', 7606-7613, 'X', 7615-7622, 'X', 7624-7631, 'X', 7633-7640, 'X', 7642-7649, 'X', 7651-7658, 'X', 7660-7667, 'X', 7669-7676, 'X', 7678-7685, 'X', 7687-7694, 'X', 7696-7703, 'X', 7705-7712, 'X', 7714-7721, 'X', 7723-7730, 'X', 7732-7739, 'X', 7741-7748, 'X', 7750-7757, 'X', 7759-7766, 'X', 7768-7775, 'X', 7777-7784, 'X', 7786-7793, 'X', 7795-7802, 'X', 7804-7811, 'X', 7813-7820, 'X', 7822-7829, 'X', 7831-7838, 'X', 7840-7847, 'X', 7849-7856, 'X', 7858-7865, 'X', 7867-7874, 'X', 7876-7883, 'X', 7885-7892, 'X', 7894-7901, 'X', 7903-7910, 'X', 7912-7919, 'X', 7921-7928, 'X', 7930-7937, 'X', 7939-7946, 'X', 7948-7955, 'X', 7957-7964, 'X', 7966-7973, 'X', 7975-7982, 'X', 7984-7991, 'X', 7993-8000, 'X', 8002-8009, 'X', 8011-8018, 'X', 8020-8027, 'X', 8029-8036, 'X', 8038-8045, 'X', 8047-8054, 'X', 8056-8063, 'X', 8065-8072, 'X', 8074-8081, 'X', 8083-8090, 'X', 8092-8099, 'X', 8099-8106, 'X', 8108-8115, 'X', 8117-8124, 'X', 8126-8133, 'X', 8135-8142, 'X', 8144-8151, 'X', 8153-8160, 'X', 8162-8169, 'X', 8171-8178, 'X', 8180-8187, 'X', 8189-8196, 'X', 8200-8207, 'X', 8209-8216, 'X', 8218-8225, 'X', 8227-8234, 'X', 8236-8243, 'X', 8245-8252, 'X', 8254-8261, 'X', 8263-8270, 'X', 8272-8279, 'X', 8281-8288, 'X', 8290-8297, 'X', 8299-8306, 'X', 8308-8315, 'X', 8317-8324, 'X', 8326-8333, 'X', 8335-8342, 'X', 8344-8351, 'X', 8353-8360, 'X', 8362-8369, 'X', 8371-8378, 'X', 8380-8387, 'X', 8389-8396, 'X', 8400-8407, 'X', 8409-8416, 'X', 8418-8425, 'X', 8427-8434, 'X', 8436-8443, 'X', 8445-8452, 'X', 8454-8461, 'X', 8463-8470, 'X', 8472-8479, 'X', 8481-8488, 'X', 8490-8497, 'X', 8499-8506, 'X', 8508-8515, 'X', 8517-8524, 'X', 8526-8533, 'X', 8535-8542, 'X', 8544-8551, 'X', 8553-8560, 'X', 8562-8569, 'X', 8571-8578, 'X', 8580-8587, 'X', 8589-8596, 'X', 8600-8607, 'X', 8609-86

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2005, 09:54:24 ; Search time 64.4163 Seconds  
(without alignments)  
826.751 Million cell updates/sec

Title: US-10-674-755-20

Perfect score: 514

Sequence: 1 LAKQTELEKLLDLDPEK.....TQKELDAALNELGPDGDEE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 498   | 96.9        | 228    | 2     | Q9L5B8 streptococc |
| 2          | 498   | 96.9        | 235    | 2     | Q9L5B2 streptococc |
| 3          | 498   | 96.9        | 249    | 2     | Q9L5D4 streptococc |
| 4          | 498   | 96.9        | 252    | 2     | Q9L583 streptococc |
| 5          | 498   | 96.9        | 360    | 2     | Q8KQK3 streptococc |
| 6          | 498   | 96.9        | 429    | 2     | Q9LAX7 streptococc |
| 7          | 498   | 96.9        | 526    | 2     | Q9LAX9 streptococc |
| 8          | 498   | 96.9        | 608    | 2     | Q8VQ55 streptococc |
| 9          | 498   | 96.9        | 744    | 2     | Q97T39 streptococc |
| 10         | 493   | 95.9        | 249    | 2     | Q9L5B7 streptococc |
| 11         | 484   | 94.2        | 211    | 2     | Q9L579 streptococc |
| 12         | 484   | 94.2        | 241    | 2     | Q9L580 streptococc |
| 13         | 481   | 93.6        | 249    | 2     | Q9L585 streptococc |
| 14         | 481   | 93.6        | 256    | 2     | Q9L590 streptococc |
| 15         | 480   | 93.4        | 502    | 2     | Q9LAX8 streptococc |
| 16         | 478   | 93.0        | 242    | 2     | Q9L562 streptococc |
| 17         | 474   | 92.2        | 209    | 2     | Q9L593 streptococc |
| 18         | 456   | 88.7        | 222    | 2     | Q9L584 streptococc |
| 19         | 324   | 63.0        | 246    | 2     | Q9L5B4 streptococc |
| 20         | 320   | 62.3        | 479    | 2     | Q9LAX2 streptococc |
| 21         | 320   | 62.3        | 481    | 2     | Q9LAX5 streptococc |
| 22         | 319   | 62.1        | 107    | 2     | Q8KQK2 streptococc |
| 23         | 313   | 60.9        | 653    | 2     | Q34097 streptococc |
| 24         | 308   | 59.9        | 213    | 2     | Q8GNS7 streptococc |
| 25         | 295   | 57.4        | 480    | 2     | Q9LAX3 streptococc |
| 26         | 250.5 | 48.7        | 211    | 2     | Q8GNT0 streptococc |
| 27         | 250.5 | 48.7        | 257    | 2     | Q9L594 streptococc |
| 28         | 243.5 | 47.4        | 227    | 2     | Q9KGS0 streptococc |
| 29         | 243.5 | 47.4        | 256    | 2     | Q9L595 streptococc |
| 30         | 243.5 | 47.4        | 461    | 2     | Q9LAX6 streptococc |
| 31         | 139   | 27.0        | 417    | 2     | Q9LAY3 streptococc |

32 130.5 25.4 225 2 Q9L591 streptococc  
33 130 25.3 415 2 Q9LAY1 streptococc  
34 129.5 25.2 222 2 Q9L577 streptococc  
35 129.5 25.2 262 2 Q9L576 streptococc  
36 129.5 25.2 415 2 Q9LAY7 streptococc  
37 128.5 25.0 246 2 Q9L578 streptococc  
38 128.5 25.0 416 2 Q9LAY8 streptococc  
39 128 24.9 619 2 Q54972 streptococc  
40 128 24.9 619 2 Q8DR10 streptococc  
41 127.5 24.8 237 2 Q9L592 streptococc  
42 127.5 24.8 395 2 Q9LAY9 streptococc  
43 127 24.7 249 2 Q9L575 streptococc  
44 126.5 24.6 255 2 Q9L581 streptococc  
45 126.5 24.6 255 2 Q9L5B6 streptococc

#### ALIGNMENTS

##### RESULT 1

Q9L5B8 PRELIMINARY; PRT; 228 AA.  
ID Q9L5B8  
AC Q9L5B8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PepA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=60;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Packham R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT disseminated clones.";  
RL J. Clin. Microbiol. 38:3663-3669(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=60;  
RA Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF253404; AAF67352.1; -  
DR InterPro; IPR009053; Prefoldin.  
FT NON\_TER 1 1  
FT NON\_TER 228 228  
SQ SEQUENCE 228 AA; 24430 MW; E6EAA953EC54EA0F CRC64;

Query Match 96.9%; Score 498; DB 2; Length 228;  
Best Local Similarity 98.1%; Pred. No. 6.9e-25;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKQTELEKLLDLDPEKTDQLDKEAEAEELDKADELPKNKADLEKEISNLEILLG 60  
|||||

Db 46 LAKQTELEKLLDLDPEKTDQLDKEAEAEELDKADELPKNKADLEKEISNLEILLG 105  
|||||

QY 61 GADSEDDTAALPNKATKKAELKTKQKELDAALNELGPDGDEE 104  
|||||

Db 106 GADSEDDTAALQNKATKKAELKTKQKELDAALNELGPDGDEE 149  
|||||

##### RESULT 2

Q9L582 PRELIMINARY; PRT; 235 AA.  
ID Q9L582  
AC Q9L582;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PepA (Fragment).

GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones.";  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3;  
RX Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF255545; AAF68098.1; -.  
FT NON\_TER 1  
FT NON\_TER 235  
SQ SEQUENCE 235 AA; 25424 MW; BFFB48C52CA8380 CRC64;

Query Match 96.9%; Score 498; DB 2; Length 235;  
Best Local Similarity 98.1%; Pred. No. 7.1e-25;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 60  
Db 56 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 115

Qy 61 GADSEDDTAALPNKLTAKKAELEKTOKELDAALNELGPDGDEE 104  
Db 116 GADSEDDTAALQNKLTAKKAELEKTOKELDAALNELGPDGDEE 159

RESULT 3  
Q9L5D4  
ID Q9L5D4 PRELIMINARY; PRT; 249 AA.  
AC Q9L5D4;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE pspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SP195;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones.";  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SP195;  
RX Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF252286; AAF69499.1; -.  
FT HSSP; P04268; 1IC2.  
FT NON\_TER 1  
FT NON\_TER 249  
SQ SEQUENCE 249 AA; 26821 MW; F8EA39225CF8D43F CRC64;

Query Match 96.9%; Score 498; DB 2; Length 249;  
Best Local Similarity 98.1%; Pred. No. 7.5e-25;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 60  
Db 67 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 126  
Qy 61 GADSEDDTAALPNKLTAKKAELEKTOKELDAALNELGPDGDEE 104  
Db 127 GADSEDDTAALQNKLTAKKAELEKTOKELDAALNELGPDGDEE 170

## RESULT 4

Q9L583  
ID Q9L583 PRELIMINARY; PRT; 252 AA.  
AC Q9L583;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE pspA (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=127;  
RX MEDLINE=20472698; PubMed=11015380;  
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;  
RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones.";  
RL J. Clin. Microbiol. 38:3663-3669 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=127;  
RX Beall B.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF255544; AAF68097.1; -.  
DR HSSP; P04268; 1IC2.  
FT NON\_TER 1  
FT NON\_TER 252  
SQ SEQUENCE 252 AA; 27260 MW; 82DEL13741F369CA2 CRC64;

Query Match 96.9%; Score 498; DB 2; Length 252;  
Best Local Similarity 98.1%; Pred. No. 7.6e-25;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 60  
Db 73 LAKKQTELEKLLSLDPEGKTQDELKAEAEALDKKADLPNKVADLEKEISNLEILLG 132

Qy 61 GADSEDDTAALPNKLTAKKAELEKTOKELDAALNELGPDGDEE 104  
Db 133 GADSEDDTAALQNKLTAKKAELEKTOKELDAALNELGPDGDEE 176

## RESULT 5

Q8KQK3  
ID Q8KQK3 PRELIMINARY; PRT; 360 AA.  
AC Q8KQK3;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Pneumococcal surface protein A (Fragment).  
GN Name=pspA;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=259/98;  
RX MEDLINE=22170754; PubMed=12183557;

```

RX DOI=10.1128/IAI.70.9.5086-5090.2002;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PspA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AY082389; AAL92494.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 360
SQ SEQUENCE 360 AA; 39575 MW; 0C09A791547A47EC CRC64;

Query Match 96.9%; Score 498; DB 2; Length 360;
Best Local Similarity 98.1%; Pred. No. 1.1e-24;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 60
DB 232 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 291

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
DB 292 GADSEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 335

RESULT 6
Q9LAX7 PRELIMINARY; PRT; 429 AA.
ID Q9LAX7
AC Q9LAX7
RC STRAIN=AC122;
RX MEDLINE=20448953; PubMed=10992499;
RA DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071818; AAF27714.1; -.
DR InterPro; IPR009053; Prefoldin.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 429
FT NON_TER 429
SQ SEQUENCE 429 AA; 47048 MW; BC1D74BBA54DA9D6 CRC64;

Query Match 96.9%; Score 498; DB 2; Length 429;
Best Local Similarity 98.1%; Pred. No. 1.3e-24;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 60
DB 254 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 313

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
DB 314 GADSEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 357

RESULT 7
Q9LAX9 PRELIMINARY; PRT; 526 AA.
ID Q9LAX9
AC Q9LAX9

```

```

DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF3296;
RX MEDLINE=20448953; PubMed=10992499;
RX DOI=10.1128/IAI.68.10.5889-5900.2000;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PspA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071816; AAF27712.1; -.
DR HSSP; P04268; IIC2.
DR InterPro; IPR011047; Quin_alc_DH_like.
FT NON_TER 526
FT NON_TER 526
SQ SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AE CRC64;

Query Match 96.9%; Score 498; DB 2; Length 526;
Best Local Similarity 98.1%; Pred. No. 1.5e-24;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 60
DB 346 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 405

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
DB 406 GADSEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 449

RESULT 8
Q8VQ55 PRELIMINARY; PRT; 608 AA.
ID Q8VQ55
AC Q8VQ55;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNIH1156;
RC Lee K.J., Bae S.M., Chung K.S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF460993; AAL67804.1; -.
DR HSSP; P06653; 1HCX.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
FT NON_TER 608
FT NON_TER 608
SQ SEQUENCE 608 AA; 67918 MW; 15F71BD62E297526 CRC64;

Query Match 96.9%; Score 498; DB 2; Length 608;
Best Local Similarity 98.1%; Pred. No. 1.8e-24;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 60
DB 222 LAKKQTELEKLLDSDPEGKTQDELKAEAEELDKKADLPNKVADLEKEISNLEILG 281

QY 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104
DB 61 GADSEDDTAALQNKLTAKKAELEKTKQKELDAALNELGPDGDEEE 104

```

Db 282 GADSEDDTAALQNKLATKKAELKTKQKELDAALNELGPDGDEE 325

## RESULT 9

```
Q97T39          PRELIMINARY;      PRT;    744 AA.
AC Q97T39;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pneumococcal surface protein A.
GN OrderedLocNames=Sp0117;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506 (2001).
DR EMBL; AE007328; AAK74303.1; -.
DR PIR; F95013; F95013.
DR HSP; P06653; 1HCX.
DR TIGR; SP0117; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR002345; Lipocalin.
DR Pfam; PF01473; CW binding 1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
KW Complete proteome.
SQ SEQUENCE 744 AA; 82764 MW; 20E5E8E7911EFD5 CRC64;
```

Query Match 96.9%; Score 498; DB 2; Length 744;  
Best Local Similarity 98.1%; Pred. No. 2.2e-24;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60  
Db 346 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 405

Qy 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEE 104  
Db 406 GADSEDDTAALQNKLATKKAELKTKQKELDAALNELGPDGDEE 449

## RESULT 10

```
Q9L5B7          PRELIMINARY;      PRT;    249 AA.
AC Q9L5B7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
```

RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones.";  
RL J. Clin. Microbiol. 38:3663-3669 (2000).

RN [2]  
RP SEQUENCE FROM N.A.

```
RC STRAIN=50;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253405; AAF67353.1; -.
FT NON_TER 1 249
FT NON_TER 231 1
SQ SEQUENCE 249 AA; 27271 MW; B4106707EF108A0B CRC64;
```

Query Match 95.9%; Score 493; DB 2; Length 249;  
Best Local Similarity 97.1%; Pred. No. 1.6e-24;  
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60  
Db 103 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 162

Qy 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEE 104  
Db 163 GADPEDDTAALQNKLATKKAELKTKQKELDAALNELGPDGDEE 206

## RESULT 11

```
Q9L579          PRELIMINARY;      PRT;    231 AA.
AC Q9L579;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
```

RT "Pneumococcal pspA sequence types of prevalent multiresistant  
RT pneumococcal strains in the United States and of internationally  
RT disseminated clones.";  
RL J. Clin. Microbiol. 38:3663-3669 (2000).

RN [2]  
RP SEQUENCE FROM N.A.

```
RC STRAIN=20;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255548; AAF68101.1; -.
DR HSP; P04268; 1IC2.
FT NON_TER 1 231
FT NON_TER 231 231
SQ SEQUENCE 231 AA; 24990 MW; A7731F3A46460186 CRC64;
```

Query Match 94.2%; Score 484; DB 2; Length 231;  
Best Local Similarity 95.3%; Pred. No. 5.6e-24;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LAKQTELEKLLDSLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 60  
Db 76 LAKQTELEKLLDNLDPEGKTQDELKAEAEALDKADELPNKVADLEKEISNLEILLG 135

Qy 61 GADSEDDTAALPNKLTAKKAELEKTKQKELDAALNELGPDGDEE 104  
Db 136 GADPEDDTAALQNKLATKKAELKTKQKELDAALNELGPDGDEE 179

```

RESULT 12
ID Q9L580 PRELIMINARY; PRT; 241 AA.
AC Q9L580;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Cherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255547; AAF68100.1; -.
DR HSSP; P04268; IIC2.
FT NON_TER 1 1
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26038 MW; BB87E1A4C25FA669 CRC64;

Query Match 94.2%; Score 481; DB 2; Length 241;
Best Local Similarity 95.2%; Pred. No. 5.8e-24;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 60
Db 77 LAKKQTELEKLLDNLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 136

QY 61 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPGDDEE 104
Db 137 GADPEDDTAALQNKLTATKAELEKTKQKELDAALNELGPGDDEE 180

RESULT 13
Q9L585 PRELIMINARY; PRT; 249 AA.
AC Q9L585;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Cherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255547; AAF68100.1; -.
DR HSSP; P04268; IIC2.
FT NON_TER 1 1
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 27050 MW; DF4D2ED9265986FA CRC64;

Query Match 93.6%; Score 481; DB 2; Length 249;
Best Local Similarity 95.2%; Pred. No. 9.4e-24;
Matches 99; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 60
Db 68 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 127

QY 61 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPGDDEE 104
Db 126 GADPEDDTAALQNKLTATKAELEKTKQKELDAALNELGPGDDEE 171

RESULT 14
Q9L590 PRELIMINARY; PRT; 256 AA.
AC Q9L590;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).
GN Name=pspA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP193;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Cherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspA sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP193;
RA Beall B.W.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254259; AAF68094.1; -.
DR HSSP; P04268; IIC2.
FT NON_TER 1 1
FT NON_TER 256 256
SQ SEQUENCE 256 AA; 27738 MW; 7F05351559AD9238 CRC64;

Query Match 93.6%; Score 481; DB 2; Length 256;
Best Local Similarity 95.2%; Pred. No. 9.6e-24;
Matches 99; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 60
Db 66 LAKKQTELEKLLDLSLDPGKTDQELDKAEAEALDKADELPNKVADLKEISNLEILG 125

QY 61 GADSEDDTAALPNKLTAKAELEKTKQKELDAALNELGPGDDEE 104
Db 126 GADPEDDTAALQNKLTATKAELEKTKQKELDAALNELGPGDDEE 169

RESULT 15
Q9LAX8 PRELIMINARY; PRT; 502 AA.
AC Q9LAX8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PspA (Fragment).

```

GN Name=pspa;  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BG8090;  
RX MEDLINE=20448953; PubMed=10992499;  
RX DOI=10.1128/IAI.68.10.5889-5900.2000;  
RA Hollingshead S.K., Becker R., Briles D.E.;  
RT "Diversity of Pspa: mosaic genes and evidence for past recombination  
RT in Streptococcus pneumoniae.";  
RL Infect. Immun. 68:5889-5900(2000).  
DR EMBL; AF071817; AAF27713.1; -.  
DR HSSP; O15813; 1D7M.  
DR InterPro; IPR011047; Quin\_alc\_DH\_like.  
DR InterPro; IPR000533; Tropomyosin.  
DR PRINTS; PR00194; TROPOMYOSIN.  
FT NON TER 502  
SQ SEQUENCE 502 AA; 55018 MW; 4E73D477CAE79B40 CRC64;  
  
Query Match 93.4%; Score 480; DB 2; Length 502;  
Best Local Similarity 94.2%; Pred. No. 2.le-23;  
Matches 98; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 1 LAKKQTELEKLLDSDPEGKTQDELDKAEAEALDKKADLPNKVADLEKEISNLEILLG 60  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
347 LAKKQTELEKLLDNDPEGKTQDELDKAEAEALDKKADLPNKVADLEKEISNLEILLG 406  
  
Qy 61 GADSEDDTAALPNKATKKALEKTKQKELDAALNELGPDGDEE 104  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
407 GADPEDDTAALQNKATKKALEKTKQKELDAALNELGPDGDEE 450

Search completed: June 21, 2005, 10:22:14  
Job time : 64.4163 secs